Unclassified

AD-A175 058

PAGE

1a REPORT SECURITY CLASSIFICA		1b MARKINGS
Unclassified		
2a SECURITY CLASSIFICATION AUTHORITY		3 DISTRIBUTION/AVAILABILITY OF REPORT
2b DECLASSIFICATION/DOWNGRADING SCHEDULE		Approved for public release; distribution unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S)		5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR·TR· 86-2173
6a NAME OF PERFORMING ORGANIZATION Honeywell Systems & Research Center	6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c ADDRESS (City, State, and ZIP Code) 3660 Marshall Street NE Minneapolis, MN 55418		7c ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332
8a NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Office of Scientific Research	8b OFFICE SYMBOL (If applicable) AFSC NM	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-84-C-0088
8c ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332		10 SOURCE OF FUNDING NUMBERS

JEC 1.2 1986

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61102F	2304	A1	

11 TITLE (Include Security Classification)

Robust Control of Multivariable and Large Scale Systems

12 PERSONAL AUTHOR(S)

John C. Doyle, Cheng Chih Chu

13a TYPE OF REPORT Final Report	13b TIME COVERED FROM Jul 84 TO Oct 85	14 DATE OF REPORT (Year, Month, Day) 14 March 86	15 PAGE COUNT 188

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Robustness, analysis, synthesis, structured singular value, Internal stability, linear fractional transformation, Coprime factorization, H∞ Optimal Control, Hankel operators, General distance problems.(continued)

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

New methods are presented for analysis and synthesis of control systems subject to structured uncertainties. The technical approach involves the structured singular value, μ, as an analysis tool and H∞ as a synthesis tool.

Alternative formulations are compared with the H∞ approach, extensions of μ to handle real parameter problems are presented, and the issue of the convergence of μ-synthesis to a global optimum is studied.

A comprehensive solution for the synthesis of general optimal controllers is given for linear lumped time-invariant systems.

The existence of an optimal solution for H∞ optimization is proved, and some properties of this solution are discovered. The method called γ·iteration is presented and its convergence properties are established.

A new algorithm for solving a class of algebraic riccati equations is obtained. Explicit error bounds for model reduction in the synthesis process are derived.

The constructions use standard matrix operations and linear algebra applied to state-space representations of linear systems.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT		21 ABSTRACT SECURITY CLASSIFICATION
X UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS		Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL James Crowley	22b TELEPHONE (Include Area Code) 767-5025	22c OFFICE SYMBOL NM

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted   SECURITY CLASSIFICATION OF THIS PAGE
All other editions are obsolete

Unclassified

OTIC FILE COPY

18. (Concluded)

γ-iteration, spectral factorization, Riccati Equations.

HC 68 BACK

TABLE OF CONTENTS

Page

NOTATION AND ABBREVIATIONS

Notation

R	the real numbers.
C	the complex numbers.
$\cdot$	$\equiv \dfrac{d}{dt}$.
$\delta(t)$	Dirac delta function.
E	expectation operator.
A^{-1}	inverse of the matrix A .
$A^{\dagger}$	pseudo-inverse of the matrix A .
$A_{\perp}$	orthogonal complement of the matrix A .
A^{T}	transpose of the matrix A .
A^{*}	complex conjugate transpose of the matrix A .
$A > 0$	the matrix A is positive definite.
$A \geq 0$	the matrix A is positive semi-definite.
$\lambda_i(A)$	the ith eigenvalue of the matrix A .
$\sigma_i(A)$	the ith singular value of the matrix A .
$\bar{\sigma}(A)$	the maximum singular value of the matrix A .
$Ker(A)$	the kernel of the matrix A when viewed as a linear operator.
$Tr[\,A\,]$	the trace of the matrix A .
$||A||_F$	Frobenius norm of the matrix A .
R_p	proper, real-rational functions.
$G^{*}(s)$	$\triangleq G^{T}(-s)$.

$j\omega$ point on the imaginary axis (real parameter $\omega \in (-\infty,\infty)$).

$j\mathbf{R}$ the imaginary axis.

T the unit circle ($|z| = 1$).

L_2 Hilbert space of matrix-valued functions which are square integrable on $j\mathbf{R}$ (or T) in the sense of inner product on $j\mathbf{R}$

$$<F,G> \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} Tr\left[F(j\omega)^*G(j\omega)\right] d\omega \quad \text{if } F, G \in L_2(j\mathbf{R})$$

(or on T

$$<F,G> \triangleq \frac{1}{2\pi} \int_{0}^{2\pi} Tr\left[F(e^{j\theta})^*G(e^{j\theta})\right] d\theta \quad \text{if } F, G \in L_2(T)).$$

H_2 the functions in L_2 which are analytic in the open right half plane (or in the unit disc) and satisfying

$$\sup_{\sigma>0} \int_{-\infty}^{\infty} Tr\left[G(\sigma + j\omega)^*G(\sigma + j\omega)\right] d\omega < \infty$$

(or

$$\sup_{0\leq r<1} \int_{0}^{2\pi} Tr\left[G(re^{j\theta})^*G(re^{j\theta})\right] d\theta < \infty).$$

$H_2^{\perp}$ the orthogonal complement of H_2 in L_2.

$\mathbf{P}_{H_2}$ the orthogonal projection from L_2 to H_2.

$\mathbf{P}_{H_2^{\perp}}$ the orthogonal projection from L_2 to $H_2^{\perp}$.

L_{∞} Banach space of matrix-valued functions which are (essentially) bounded on $j\mathbf{R}$ (or T).

H_{∞} the functions in L_{∞} with with a bounded analytic continuation to the right half plane (or inside the unit disc).

$||G||_2$ L_2/H_2 norm

$$\triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} Tr\left[G(j\omega)^*G(j\omega)\right] d\omega \quad \text{if } G \in L_2(j\mathbf{R}) \text{ or } H_2(j\mathbf{R}).$$

$$\triangleq \frac{1}{2\pi} \int_{0}^{2\pi} Tr\left[G(e^{j\theta})^*G(e^{j\theta})\right] d\theta \quad \text{if } G \in L_2(T) \text{ or } H_2(T).$$

$||G||_\infty$ $\quad$ L_∞/H_∞ norm

$$\triangleq \operatorname*{ess\ sup}_{\omega \in \mathbf{R}} \bar{\sigma}\left[G(j\omega)\right] \qquad \text{if } G \in L_\infty(j\mathbf{R}) \text{ or } H_\infty(j\mathbf{R}).$$

$$\triangleq \operatorname*{ess\ sup}_{\iota \in [0,2\pi)} \bar{\sigma}\left[G(e^{j\iota})\right] \qquad \text{if } G \in L_\infty(T) \text{ or } H_\infty(T).$$

$||M||$ $\qquad$ the operator norm of M.

$\mathbf{M}_G$ $\qquad$ multiplicative (Laurent) operator generated by $G \in L_\infty$.

$\mathbf{H}_G$ $\qquad$ *Hankel* operator (matrix) generated by $G \in L_\infty$

$\mathbf{T}_G$ $\qquad$ *Toeplitz* operator (matrix) generated by $G \in L_\infty$.

$\mathbf{F}_l(P,K)$ $\qquad$ linear fractional transformation of P and K.

$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ $\quad \triangleq D + C(sI-A)^{-1}B$ (shorthand notation transfer functions).

When R is used as a prefix, it denotes real-rational. The superscripts "n" and "$p \times m$" (as in $\mathbf{R}^n$ and $\mathbf{R}^{p \times m}$) will denote the corresponding n-vectors and $p \times m$ matrices.

Abbreviations

rhp	right half plane of complex number plane.
lhp	left half plane of complex number plane.
rcf	right coprime factorization.
lcf	left coprime factorization.
SISO	single-input/single-output.
MIMO	multiple-input/multiple-output.
LQG	linear-quadratic-Gaussian.
WHKB	Wiener-Hopf-Kalman-Bucy.
LFT	linear fractional transformation.
GEP	generalized eigenvalue problem.
IOF	inner-outer factorization.
CIF	complementary inner factor.
ARE	algebraic Riccati equation.
GDP	general distance problem.

CHAPTER 1

INTRODUCTION

This report presents some new methods for analysis and synthesis of control systems for robust performance in the presence of structured uncertainty. It builds on the results of Doyle (1984). The technical approach involves the structured singular value, μ, as an analysis tool and H_∞ as a synthesis tool. These are combined to form the basis for μ-synthesis.

The major contributions of this report with their corresponding chapters are:

(1) A comparison is given of H_∞ performance and robustness formulations with some alternatives using other norms. Performance for bounded magnitude time signals are found in terms of H_∞ norms on transfer functions. (1)

(2) Extensions of μ to handle real parameter variations are considered. Improved bounds are obtained for this problem. (1)

(3) The issue of convergence of μ-synthesis to a global optimum is studied. It is shown that the global solution to the μ-synthesis problem can be found in the constant (or equivalently acausal) case. This provides useful information for the general case and is encouraging regarding the prospects of obtaining similar results there. (1)

(4) A comprehensive and unified treatment is given to the synthesis of general optimal controllers for linear lumped time-invariant systems. (2)

(5) The existence of an optimal solution for the general H_∞-optimization formulation is proven. In particular, it is shown that there exists a real-rational optimal solution when the original data is real-rational. (3)

(6) An iterative procedure, called γ-iteration, is discussed which reduces the general distance formulation to an equivalent solvable best approximation formulation. Some tight and comput-

able bounds are derived and properties which guarantee rapid convergence of the iteration are established. (4)

(7)   A new and efficient algorithm for solving a class of algebraic Riccati equations that arises in the γ-iteration is obtained. (5)

(8)   State-space representations and reliable algorithms are derived in parallel to the development of the other theory throughout this work. (5)

(9)   Explicit error bounds are obtained which make model reduction in the synthesis process possible and simplify the complexity of the controller significantly. (6)

Based on the results in this report and in Doyle(1984), an experimental software package has been developed which uses only standard matrix operations and linear algebra techniques. Several example designs have been performed which have been very successful and most encouraging. These example designs will be presented elsewhere. Many of the results presented in this report have appeared in papers by the authors.

1.1 General Analysis Framework

This chapter will review some basic methods for analyzing the performance and robustness properties of feedback systems. The particular approach taken here is from [D14]-[D19] and [C2]-[C4] which builds on results by many other researchers. In this context, analysis refers to the process of determining whether a system with a given controller has desired characteristics, whereas synthesis refers to the process of finding a controller that gives desired characteristics, usually expressed in terms of some analysis method. This is the fairly standard usage of these terms in the control community. It should be obvious that the question of analysis must be settled to some degree before a reasonable synthesis problem can be posed. The formal analysis and synthesis techniques discussed in this report are only some of the methods that might make up the overall process of engineering design.

The general framework to be used in this report is illustrated in the diagram in Figure 1.1. Any linear interconnection of inputs, outputs, commands, perturbations, and a controller can be rearranged to match this diagram. For the purpose of analysis the controller may be thought of as just another system component and the diagram reduces to that in Figure 1.2. The analysis problem involves determining whether the error e remains in a desired set for sets of inputs v and perturbations Δ. Analysis methods differ on the description of these sets and the assumptions on the interconnection structure G. In this report G will be taken to be a linear, time-invariant, lumped system and be represented by a rational transfer function. The convolution kernel associated with G will be denoted as g, so G is a real-rational matrix function of a complex variable and g is a matrix function of time. The interconnection structure G can be partitioned so that the transfer function from v to e can be expressed as the linear fractional transformation

$$e = F_u(G, \Delta) v = \left[G_{22} + G_{21}\Delta(I - G_{11}\Delta)^{-1} G_{12} \right] v.$$

The external input v is an additive signal entering the system and is typically used to model disturbances, commands, and noise. The alternative descriptions of the sets to which v is assumed to belong and the corresponding performance requirements on e will be considered in the next section. It is generally inadequate in modeling systems for control design to consider uncertainty only in the form of uncertain additive signals [H4]. The system model itself typically has uncertainty which can have a significant impact on system performance. This uncertainty is a consequence of unmodeled dynamics and parameter variations and is modeled as the perturbations Δ to the nominal interconnection structure G. Note that the uncertainty

modeled as Δ has a very different effect from that of v on the performance of the system. For example, perturbations can cause a nominally stable system to become unstable, which v cannot do. Techniques for modeling plant perturbations and analyzing their impact on stability will be considered in Section 3.

At the heart of any theory about control are the assumptions made about G, v and Δ, as well as the performance specifications on e. These assumptions determine the analysis methods which can be applied to obtain conclusions about system performance. A desirable objective is to make weak assumptions but still arrive at strong conclusions and the inevitable tradeoff implied by this objective drives the development of new theory. The control theoreticians role may be viewed as one of developing methods that allow the control engineer to make assumptions which seem relatively natural and physically motivated. The ultimate question of the applicability of any mathematical technique to a specific physical problem will always require a "leap of faith" on the part of the engineer and the theoretician can only hope to make this leap smaller.

The methods in this chapter provide a powerful set of techniques for modeling and analyzing uncertain systems. To provide a context for these methods consider an alternative analysis technique based on simulation. Modern simulation methods allow for a wide variety of plant models and performance specifications and as a consequence are an essential part of the control engineers toolkit. The price paid for this flexibility is that the system inputs and perturbations can be considered only one at a time. This chapter focuses on describing *sets* of signals and perturbations and drawing conclusions on worst-case performance for entire sets at once. An important issue to be addressed here is the description of sets which adequately model physical phenomena. It can be just as limiting to have models which allow signals or perturbations which have no physical motivation but severely impact performance (of the model) as it is to have models that ignore critical physical phenomena. The methods in this chapter are aimed at allowing for signal and perturbation models that naturally match the physical phenomena that the models are intended to represent. The price paid for *this* flexibility, in contrast with simulation, is that very restrictive assumptions , in particular linearity, must be made about the nominal interconnection structure. Nonlinearities can only be handled as perturbations.

It is beyond the scope of this report to give a thorough discussion of the relationship between models and the physical systems they represent. Attention will be given to the main assumptions that have been proven useful in practice, and some comparison of the alternatives. The particular focus of this chapter is

on developing techniques that allow very precise analysis of systems which have fairly standard performance requirements and uncertainty models in terms of additive noise and plant perturbations. While the "best" assumptions for engineering purposes will always be a matter of debate, it is clear that for any given set of assumptions it is desirable to have very precise analysis techniques. The ideal would be necessary and sufficient conditions for the satisfaction of a performance specification in the presence of sets of inputs and perturbations. Additionally, the conditions should be computable or should at least yield bounds which give useful estimates of system performance. With such methods, the engineer can focus directly on the relationship between uncertainty assumptions and system performance without worrying about potential gaps caused by inadequate analysis techniques.

The assumptions about v and the system performance specifications on e are considered in Section 2. A basic requirement is that the nominal system $(\Delta = 0)$ be stable. Recall that this is analysis of the closed-loop system with controller in place so it is assumed that G is stable, which will be taken to mean having no rhp poles. Performance will be expressed in terms of e being contained in a specified set bounded in power, energy, or magnitude. In Section 2 only uncertainty in v will be considered so this will be referred to as *nominal performance* to indicate that $\Delta = 0$. Nominal performance will be seen to be equivalent to a norm test on G_{22}. The main focus of this report will be on $\| \bullet \|_\infty$, but other norms of practical interest will be considered and briefly compared. It will be argued that the $\| \bullet \|_\infty$ norm is a useful and flexible norm for studying performance.

Section 3 considers stability in the presence of perturbations. This will be referred to as *robust stability* with robust used here to indicate that the property of stability is maintained under perturbations. For simple unstructured perturbations, this also leads naturally to a $\| \bullet \|_\infty$ norm test, but now on G_{11}. The $\| \bullet \|_\infty$ norm thus provides a single norm which handles both nominal performance and robust stability. Unfortunately, norm bounds are inadequate in dealing with more realistic models of plant uncertainty involving structure and more complicated mathematical objects involving the structured singular value, μ, are required.

The methods outlined in Sections 2 and 3 allow for assessing either nominal performance or robust stability. Obviously, it would be desirable to treat performance with both noise and perturbations occurring simultaneously. Section 3 concludes with this problem and shows that this also leads to tests using μ, but now involving the entire transfer function G. Thus μ emerges as an essential analysis tool in dealing with

robust performance as well as with structured perturbations.

Section 4 briefly reviews H_∞ and H_2 optimal control. The H_∞ methods combine with the properties of μ discussed in Section 5 to form the basis for μ-synthesis. Chapters 2-7 of this report expand on the theory outlined in Section 4.

The mathematical properties and computation of μ are briefly taken up in Sections 5 for the case of complex perturbations and 6 for the real case. Here μ is viewed as a natural generalization of both spectral radius and spectral norm, and this viewpoint leads to useful characterizations of μ in terms of these more familiar quantities. One consequence is that estimates for μ can be obtained by scaling of ordinary singular values. The implications of this approach for synthesis are also briefly considered in Section 7.

The main results of this chapter will be expressed as theorems which are each instances of the following form of a General Analysis Theorem:

General Analysis Theorem (GAT):

$$\boxed{\text{Performance}} \quad \text{for all} \quad \boxed{\text{Uncertainty}}$$

$$\text{iff} \quad \boxed{\text{Analysis Test}}$$

As implied by the form of this "theorem", this chapter will focus on necessary and sufficient conditions for performance in the presence of uncertainty. The uncertainty will be combinations of input signals, perturbations, and parameter variations. Performance will be simply stability or stability plus a bound on the error e. It is hoped that by organizing the many alternative methods of performance and robustness analysis in this way, it will be easier to compare the assumptions and their relative merits.
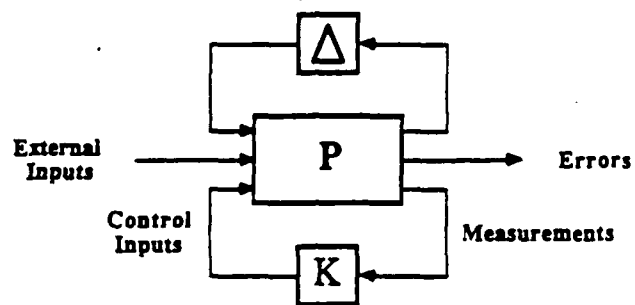
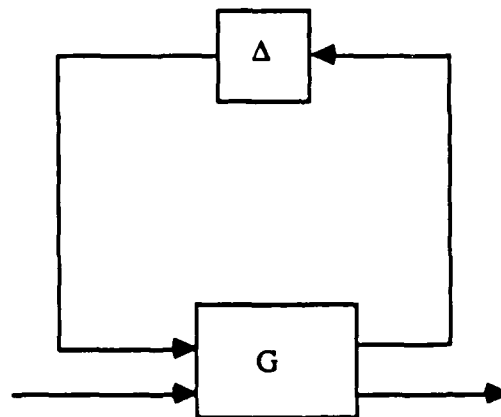Figure 1-1. General Framework for Control System Design

Figure 1-2. General Diagram with Controller Absorbed into G.

1.2 Nominal Performance

This section considers performance in terms of bounds on e in the presence of uncertain bounded inputs v. Bounds for both v and e are expressed in terms of signal power, energy, or magnitude. Such descriptions are standard within the control theory community and a rigorous treatment will not be given here. The focus is on comparing the resulting tests on G implied by the alternative assumptions. While other assumptions on signals could be considered, these are the most common both in the literature and in practice. To simplify the comparisons, assume temporarily that the signals are scalars so that spatial norms are not an issue. This will focus attention on the contrasts between the alternative descriptions of the time content of the norms, which is far more significant.

The three alternative assumptions about the signals is that they are bounded in "average power", "total energy", or magnitude. The terms power and energy are used here in a generalized sense to indicate that integrals of the square of the signals are involved. This is standard usage of these terms within the control community. Suppose that v is a function of time such that on any finite interval it is square integrable. Then we may obtain bounds on v in terms of:

(1) Power: $\quad BP \quad = \left\{ v \ \Big| \ \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |v(t)|^2 \ dt < 1 \right\}$

(2) Energy: $\quad BL_2 \quad = \left\{ v \ \Big| \ \|v\|_2^2 = \int_{-\infty}^{\infty} |v(t)|^2 \ dt < 1 \right\}$

(3) Magnitude: $BL_\infty = \left\{ v \ \Big| \ \|v\|_\infty = \operatorname*{ess\,sup}_t |v(t)| < 1 \right\}$

The prefix B denotes the unit ball. The bounds are scaled to 1 since any other scaling can simply be absorbed into the interconnection structure G. Likewise, any weighting or coloring filter can be absorbed into G so that only unweighted signals need be considered. Note that in practice, the use of weightings on both v and e are essential to reflect the varying spatial and frequency content of both the input signals and the performance specifications. For simplicity, all signals will be assumed to be complex.

For nominal performance the GAT takes the form:

GAT for Nominal Performance:

$$\boxed{e \in \text{Performance Set}} \quad \text{for all} \quad \boxed{v \in \text{Input Set}}$$

$$\text{iff} \quad \boxed{\text{Norm Test}}$$

Here the performance and input sets are taken to be either BP, BL_2, or BL_∞ as defined above with the additional input set of sinusoids also being considered. Table 1 gives the norm test for each combination that makes sense where

$$\|G\|_2 = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace} \left[G(j\omega)^* \, G(j\omega) \right] \, d\omega \right)^{1/2}$$

$$\|G\|_\infty = \sup_\omega \overline{\sigma} \left[G(j\omega) \right]$$

$$\|g\|_1 = \int_0^{\infty} |g(t)| \, dt.$$

The $\|G\|_\infty$ norm is defined for matrices since this will be used in the remaining sections.

Table 1. Performance Summary

v \ e	Power	Energy	Magnitude
Power	$\|G\|_\infty$	∞	∞
Energy	0	$\|G\|_\infty$	$\|G\|_2$
Magnitude	$\|G\|_\infty$	∞	$\|g\|_1$
Sinusoids	$\|G\|_\infty$	∞	$\|G\|_\infty$

In each case the norm test is the indicated norm being less than or equal to one; the norm would be applied to G_{22} (the subscript is dropped in the table). For example, the upper left hand corner case would yield the following version of the GAT:

Performance Theorem (Power/Power):

$$e \in BP \ \forall \ v \in BP$$

$$\text{iff } \|G_{22}\|_\infty \leq 1.$$

Some of the entries in Table 1 are either 0 or ∞. The 0 occurs because for any stable G, $v \in BL_2$ yields e of zero average power. The ∞ occurs because for the indicated combinations any $G \neq 0$ allows unbounded outputs for some bounded inputs. These combinations are clearly of no practical or theoretical interest.

Note that $\|G\|_\infty$ is the clearly the most common norm in Table 1 and will be the norm used in the rest of this paper. It results from several different assumptions of practical and theoretical interest. Perhaps the most useful are $v \in BP$ and $e \in BP$ and the cases of sinusoidal inputs. The vector case is no different provided the spatial part of the vector norm in each case is taken to be the usual Euclidean 2-norm. The $\|G\|_2$ and $\|g\|_1$ norms each appear only once. The $\|G\|_2$ appears for $v \in BL_2$ and $e \in BL_\infty$. While this appears to be a strange combination, it could have some significance in some problems. To my knowledge, no one has described control problems in this way.

The standard assumptions which lead to $\|G\|_2$ are not covered in the table but concern situations where either v has a fixed power spectrum and $e \in BP$ or v is a stochastic process with fixed power spectral density function and performance is measured in terms of the variance of e. The $\|G\|_2$ also arises when v is a fixed signal and $e \in BL_2$. Zames and others ([Z2],[Z3]) have argued that these assumptions are often not appropriate for control problems. At the heart of this argument is the observation that when frequency-dependent weights are use to shape the spectral content of signals and performance specifications, it may be a better model of physical reality to view inputs as being, for example, *bounded* in power than characterized by a perfectly known, fixed power spectrum. In addition, it is usually quite easy to find performance weights that turn a problem specified in terms of $\| \bullet \|_2$ into one involving only $\| \bullet \|_\infty$. While this would be of no particular value when only uncertain inputs are considered, it could prove quite useful, as will be seen, when uncertainty in the form of plant perturbations are included. While these issues still remain largely unresolved, it is clear that the popularity of $\| \bullet \|_2$ is due in great part to its convenient mathematical properties, which are substantial and well-known. Fortunately, this distinction between $\| \bullet \|_2$ and $\| \bullet \|_\infty$ is becoming less significant.

The $\|g\|_1$ norm cannot be dismissed as lightly. The assumptions which lead to it, that both v and e are in BL_∞, are very appealing. It is often the case in practice that the critical issue is the magnitude of signals and not their power or energy. In fact, it could be argued that this would be the obvious norm of choice for most engineering problems were it not for the mathematical difficulties associated with $\|g\|_1$. For example, it seems more difficult to shape the spectral content of signals in BL_∞ using weights than for signals assumed to be in BP. This is of great practical significance since it is typically critical in achieving good designs to take advantage of what is known about the frequency content of signals. In addition, $\|g\|_1$ is very difficult to work with analytically. It has no useful sets of invariants analogous to the inner or all-pass functions for $\|G\|_2$ and $\|G\|_\infty$. There is no synthesis methodology for optimizing $\|g\|_1$ except in very special cases [Pearson].

Given the difficulties associated with directly synthesizing for $\|g\|_1$ and its potential practical importance, it is interesting to ask how $\|g\|_1$ relates to $\|G\|_\infty$. In particular, it is important to know that optimizing $\|G\|_\infty$ will not do great violence to $\|g\|_1$. Since the constant term in G simply adds to $\|g\|_1$, suppose for the moment that G is strictly proper. It can be shown that

$$\|G\|_\infty \leq \|g\|_1 \leq (2n+1)\|G\|_\infty \qquad (2.1)$$

where n is the McMillan degree of G. The left hand bound is well known and the right hand side is proven at the end of this section. In fact, a stronger result is proven, namely that

$$\|g\|_1 \leq 2 \sum_i \sigma_i \qquad (2.2)$$

where $\{\sigma_i\}$ are the singular values of the Hankel operator associated with G. Given that the response of many systems can be approximated at least crudely by fairly low order systems, this bound suggests that $\|G\|_\infty$ may often be a reasonable approximation to $\|g\|_1$.

It should be noted that examples can be constructed for any n so that all the bounds in (2.1) and (2.2) are achieved. The examples that achieve the upper bounds are pathological and require G to have poles and zeros widely spread and the inputs to have very broad spectral content. This suggests that even more useful bounds may be obtained when additional assumptions such as restrictions on pole and zero locations and input signal bandwidth are imposed. This is clearly an area that will require additional research.

Proof of bound $\|g\|_1 \leq 2n\|G\|_\infty$. (the idea for this proof is due to I. Gohberg)

Let $G(s) = c(sI - A)^{-1}b$, where (A, b, c) is an internally balanced minimal realization (Moore) so that

$$\int_0^\infty e^{A't} c' c e^{At} dt = \int_0^\infty e^{At} bb' e^{A't} dt$$

$$= \sum = \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n) > 0 . \tag{2.3}$$

Let $\phi_i(t) = \frac{1}{\sqrt{\sigma_i}} e_i' e^{At} b$ and $\psi_i(t) = \frac{1}{\sqrt{\sigma_i}} c e^{At} e_i$ where e_i is the ith unit vector. Note that (2.3) implies that $\|\phi_i\|_2 = \|\psi_i\|_2 = 1$. Expanding $g(t) = c e^{At} b = \sum_i \sigma_i \psi_i(\frac{t}{2}) \phi_i(\frac{t}{2})$ yields

$$\|g\|_1 = \int_0^\infty \left| \sum_i \sigma_i \psi_i\left(\frac{t}{2}\right) \phi_i\left(\frac{t}{2}\right) \right| dt$$

$$\leq 2 \sum_i \sigma_i \int_0^\infty |\psi_i(r) \phi_i(r)| dr$$

$$\leq 2 \sum_i \sigma_i$$

Furthermore, since $\|G\|_\infty \geq \sigma_i$, $\forall i$, $\|g\|_1 \leq 2n\|G\|_\infty$.

1.3 Robust Stability and Robust Performance

In this section, we will consider plant perturbations, a type of uncertainty entirely different from uncertain input signals. Since plant perturbations can destabilize a nominally stable system, the first issue to be addressed is robust stability. Robust performance will be treated at the end of this section. In what follows, it makes no difference whether Δ is a constant complex , rational, or real-rational matrix so for simplicity it will be assumed constant complex. Stability will be taken to mean that the perturbed system has no closed rhp poles. Under these assumptions, we have the following simple and well-known theorem([Z1],[D10]):

Theorem RSU (Robust Stability, Unstructured):

$$F_u(G, \Delta) \text{ stable } \forall \; \Delta \, , \bar{\sigma}(\Delta) < 1$$

$$\text{iff } \|G_{11}\|_\infty \leq 1.$$

While the $\| \bullet \|_\infty$ norm had some reasonable competition for analyzing nominal performance, it is clearly the norm of choice for robustness analysis. While it is possible to use other norms in theorems such as the above to obtain sufficient conditions for stability, only $\| \bullet \|_\infty$ yields necessary and sufficient conditions. The only change that can be made and still have iff is to allow other spatial norms. In contrast, $\| \bullet \|_2$ norm cannot even be used to obtain sufficient conditions for robust stability.

The term unstructured refers to the fact that Δ is assumed to be bounded but otherwise unknown. Typically weights are used when modeling plant uncertainty to reflect the frequency and spatial variation of the perturbations. These weights can always be absorbed into the nominal interconnection structure so in that sense it is no loss of generality to assume a uniform norm bound on Δ. It is in the assumption that no structural information is available for Δ that limits the usefulness of Theorem RSU. In practical problems, it is generally the case that the uncertainty consists of parameter variations and *multiple* norm-bounded perturbations. Using only a single norm-bounded perturbation for analysis is rarely adequate. Parameter variations typically arise because of uncertain coefficients in differential equation models of physical systems and involve real scalars. Norm-bounded perturbations often arise when trying to capture the effect of unmodeled dynamics and are themselves dynamic systems. This would typically lead to norm-bounded real-rational perturbations, but for analysis, it is sufficient to instead consider constant complex matrix perturbations.

Any linear interconnection of inputs, outputs, transfer functions, parameter variations, and perturbations can be rearranged to fit the diagram in figure 2, where $\bar{\sigma}(\Delta) < 1$ but Δ is block-diagonal. This is an obvious consequence of the fact that composition of linear fractional transformations are linear fractional, and it holds for perturbations to transfer functions as well as elements of state-space realizations. Reducing to the uniform norm bound typically requires the absorption into the nominal interconnection structure of scalings and weights. Then Δ will be a member of a set like

$$\boldsymbol{\Delta} = \{ \text{diag} \left(\delta_1, \delta_2, \ldots, \delta_m, \Delta_1, \Delta_2, \ldots, \Delta_n \right)$$

$$\left| \; \delta_i \in \mathbf{R}, \Delta_j \in \mathbf{C}^{k_j \times k_j} \right\} \tag{3.1}$$

or its bounded subset:

$$\mathbf{B\Delta} = \left\{ \Delta \in \boldsymbol{\Delta} \; \middle| \; \bar{\sigma}(\Delta) < 1 \right\}. \tag{3.2}$$

It is possible to define more general sets involving, for example, repeated perturbations, and these will be considered in Section 5. Nonsquare perturbations can easily be handled in what follows by augmenting the interconnection structure with rows or columns of zeros. It should be noted that although the block diagonal perturbation structure with square, uniformly bounded blocks can be used without loss of generality, it may be desirable from a computational point of view to use other structures. This particular structure is chosen because it is mathematically general and conceptually elegant.

Given $\Delta \in \mathbf{B\Delta}$ Theorem RSU could be used to obtain sufficient conditions for robust stability, but the test could be arbitrarily conservative. That is, it is easy to construct examples where $\|G_{11}\|_\infty$ can be made arbitrarily large but no $\Delta \in \mathbf{B\Delta}$ leads to instability. In order to obtain a precise generalization of Theorem RSU to handle structured uncertainty, we need the structured singular value, μ [D13]. The positive real-valued function μ satisfies the property

$$\det \left(I - M\Delta \right) \neq 0 \;\; \text{for} \; \forall \; \Delta \in \; \boldsymbol{\Delta}, \bar{\sigma}(\Delta) < \gamma$$

$$\text{iff} \;\; \gamma\mu(M) \leq 1. \tag{3.3}$$

Note that μ is a function of M that depends on the structure of the Δ's in $\boldsymbol{\Delta}$. This dependency is typically not represented explicitly. If $\mu(M) \neq 0$, that is $\exists \Delta \in \boldsymbol{\Delta}$ such that $\det(I - M\Delta) = 0$, then

$$\frac{1}{\mu(M)} = \min_{\Delta \in \boldsymbol{\Delta}} \left\{ \sigma(\Delta) \; \middle| \; \det(I - M\Delta) = 0 \right\}. \tag{3.4}$$

Unfortunately, (3.4) is not typically very useful in computing μ since the implied optimization problem is cumbersome and can have multiple local maxima which are not global. Computation of μ is a complicated problem and some results will be given in Sections 5 and 6. For now, assume μ is the function defined above.

With these definitions, the correct generalization of Theorem RSU to structured uncertainty is

Theorem RSS (Robust Stability, Structured) :

$$F_u(G, \Delta) \text{ stable } \forall \Delta \in B\Delta$$

$$\text{iff } \|G_{11}\|_\mu \leq 1$$

where

$$\|G\|_\mu \overset{\text{def}}{=} \sup_\omega \mu[G(j\omega)]. \tag{3.5}$$

Note that $\|G\|_\mu$ is not actually a norm, but the notation is convenient. Note also that it depends not only on G but also the assumed structure of Δ.

The methods outlined above allow for analyzing either nominal performance or robust stability. Obviously, it would be desirable to treat performance with both noise and perturbations occuring simultaneously [D14]. The following theorem addresses exactly this problem. The proof is given at the end of this section.

Theorem RP:

$$F_u(G, \Delta) \text{ stable and } \|F_u(G, \Delta)\|_\infty < 1 \; \forall \Delta \in B\Delta$$

$$\text{iff } \|G\|_\mu \leq 1$$

where μ is taken w.r.t. the structure

$$\tilde{\Delta} = \{\tilde{\Delta} = \text{diag } (\Delta, \; \Delta_{n+1}) \; \Big| \; \Delta \in \Delta\}.$$

This theorem is the real payoff for using μ. It gives necessary and sufficient conditions for robust performance in the presence of structured uncertainty. It's made possible by the equivalence of performance and robust stability when using $\| \bullet \|_\infty$. The block Δ_{n+1} may be thought of loosely as a "performance block" used to turn the performance condition into a robust stability condition and finally into a test using μ. Note

that μ is computed for the full G and is taken with respect to an augmented structure. The analysis results presented in this paper are summarized in Table 2.

Table 2. Analysis Summary

Performance Perturbation	Stability	$e \in BP$
$\Delta = 0$	No C_+ poles	$\|G_{22}\|_\infty \leq 1$
$\bar{\sigma}(\Delta) < 1$	$\|G_{11}\|_\infty \leq 1$	$\|G\|_\mu \leq 1$
$\Delta \in \mathbf{B\Delta}$	$\|G_{11}\|_\mu \leq 1$	$\|G\|_\mu \leq 1$

Proof of Theorem RP: From Theorem RSU,

$$\|F_u(G, \Delta)\|_\infty \leq 1$$

$$\text{iff } \det(I - F_u(G, \Delta)\Delta_{n+1}) \neq 0 ,$$

$$\forall s = j\omega , \ \forall \Delta_{n+1} , \ \bar{\sigma}(\Delta_{n+1}) < 1 .$$

Similarly,

$$F_u(G, \Delta) \text{ is stable } \forall \Delta \in \mathbf{B\Delta}$$

$$\text{iff } \det(I - G_{11}\Delta) \neq 0 \ \forall s = j\omega , \ \forall \Delta \in \mathbf{B\Delta} .$$

Since

$$\det(I - G\tilde{\Delta}) = \det(I - G_u\Delta)\det(I - F_u(G, \Delta)\Delta_{n+1}),$$

the result follows immediately from definition of μ and $\| \cdot \|_\mu$.

1.4 Overview of Optimal Synthesis Theory

The previous sections showed how nominal performance and robust stability with unstructured uncertainty could be treated using the H_∞ framework. More complicated issues like structured uncertainty and robust performance require μ. This section contains a review of the general H_2 and H_∞ optimal synthesis theory using the framework depicted in Figure 1-3. The approach outlined in this section was developed in Doyle (1983,1984), Chu and Doyle (1984,1985), and Chu (1985). Although the focus of this report is on H_∞ methods, it is useful to put these newer methods in a context which includes the more familiar H_2 theory. The general approach taken in this section makes this easy since the H_2 and H_∞ theory can be developed in parallel up to the final step, called the General Distance Problem (GDP), which is the subject of the remainder of this report.

In Figure 1-3, the transfer function matrix P is the interconnection structure from the nominal model of the system and the transfer function matrix K is the "controller" to be designed. The variable v consists of all external inputs, e are the error signals which are to be regulated, u are the control inputs, and y are the measurements. This general framework covers all standard lumped linear time-invariant filtering and control problems. Attention will be restricted to the lumped case so all transfer functions will be assumed to be rational.
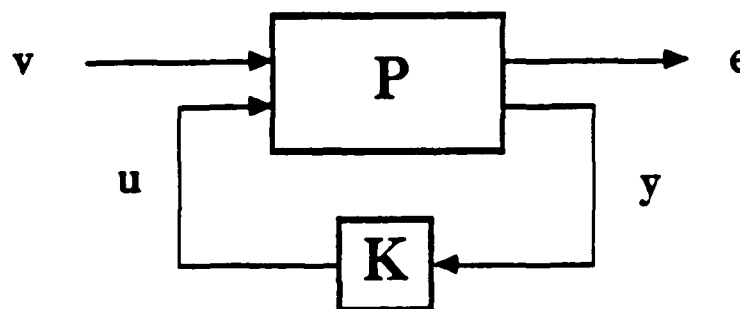

Figure 1-3. General Framework for Synthesis

Partitioning P accordingly, the closed-loop transfer function matrix can be written as the following linear fractional transformation (LFT)

$$e = \mathbf{F}_l(P,K)\, v = (\, P_{11} + P_{12}K\,(I - P_{22}K)^{-1}P_{21}\,)\, v \tag{4-1}$$

where

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in R^{(p_1 + p_2) \times (m_1 + m_2)} \quad , \quad P_{ij} \in R^{p_i \times m_j} . \tag{4-2}$$

The H_2 and H_∞ synthesis problem is one of finding a stabilizing $K \in R^{m_2 \times p_2}$ such that the performance measure

$$|| \mathbf{F}_l(P,K) ||_\alpha \qquad \text{for } \alpha = 2 \text{ or } \infty \tag{4-3}$$

is minimized. For nontriviality, assume that $p_1 > m_2$ and $m_1 > p_2$.

The first step is to find

$$K_{\bullet} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \in R^{(m_2 + p_2) \times (p_2 + m_2)} \tag{4-4}$$

such that

$$\mathbf{F}_l(\, P \; ; \mathbf{F}_l(\, K_{\bullet} \; ; Q \,)) = \mathbf{F}_l(\, T \; ; Q \,) = T_{11} + T_{12}QT_{21} \in RH_\infty^{p_1 \times m_1} \tag{4-5}$$

is stable and affine for any $Q \in RH_\infty^{m_2 \times p_2}$. This is the "Youla parametrization" of all stabilizing controllers and is obtained by finding coprime factorizations of P over the ring of stable rationals and solving a double Bezout identity to obtain the coefficients of $K_{\bullet}$ (Youla, Jabr, and Bongiorno (1976), Desoer, Liu, Murray, and Saeks (1980)). For simplicity, the superscripts denoting spatial dimensions will be suppressed in the remainder of this section. All quantities are assumed to be of compatible but otherwise arbitrary dimension.

We are interested in a particular $K_{\bullet}$ which results in T_{12} and T_{21} being inner and co-inner respectively; that is, $T_{12}^* T_{12} = I$ and $T_{21}T_{21}^* = I$. This requires a coprime factorization with inner numerator. In addition, $T_\perp$ and $\tilde{T}_\perp$ can be found so that $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}$ are square and inner. $T_\perp$ and $\tilde{T}_\perp$ are called complementary inner factors (CIF). These factorizations can be carried out using

standard real matrix operations on state-space representations (Doyle (1983,1984), Chu and Doyle (1984)). Computation of state-space realizations of K_*, T, $T_\perp$, and $\tilde{T}_\perp$ from one for P involves solving two standard algebraic Riccati equations (ARE).

Because both the $||\cdot||_2$ and $||\cdot||_\infty$ norms are invariant under multiplication by square inner matrices, an alternative expression is possible. For any $Q \in RH_\infty$,

$$|| T_{11} + T_{12}QT_{21} ||_\alpha$$

$$= \left|\left| \begin{bmatrix} -T_{12} & T_\perp \end{bmatrix}^* (T_{11} + T_{12}QT_{21}) \begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}^* \right|\right|_\alpha$$

$$= \left|\left| \begin{bmatrix} -T_{12}^* T_{11} T_{21}^* - Q & -T_{12}^* T_{11} \tilde{T}_\perp^* \\ T_\perp^* T_{11} T_{21}^* & T_\perp^* T_{11} \tilde{T}_\perp^* \end{bmatrix} \right|\right|_\alpha$$

$$= \left|\left| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right|\right|_\alpha \qquad (4\text{-}6)$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} -T_{12}^* \\ T_\perp^* \end{bmatrix} T_{11} \begin{bmatrix} T_{21}^* & \tilde{T}_\perp^* \end{bmatrix}. \qquad (4\text{-}7)$$

A state-space realization for R can be obtained from a corresponding one for P using the factorizations involved in obtaining K_* (Doyle (1984), Chu (1985)). In particular, this approach yields an R with all its poles in the open *rhp*, i.e., R is completely unstable.

Up to equation 4-6 the H_2 and H_∞ problems can be handled in parallel and the same factorization techniques can be used to reduce 4-3 to 4-6. It is in minimizing 4-6 that the two cases differ substantially. Here, the $\alpha = 2$ case is particularly simple. Since

$$\left|\left| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right|\right|_2 = \left(\left|\left| R_{11} - Q \right|\right|_2^2 + \left|\left| \begin{bmatrix} 0 & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right|\right|_2^2 \right)^{1/2} \qquad (4\text{-}8)$$

the optimal Q can be found from

$$\min_{Q \in RH_\infty} || R_{11} - Q ||_2 = || R_{11} - Q_{opt} ||_2 \qquad (4\text{-}9)$$

where

$$Q_{opt} = P_{H_2}(R_{11}) + R_{11}(\infty) = R_{11}(\infty). \tag{4-10}$$

Equations 4-9 and 4-10 follow immediately from the Hilbert space structure of L_2 and the fact that R_{11} is completely unstable. In most H_2 problems, R_{11} is strictly proper, so $Q_{opt} = 0$ and $K_{opt} = K_{11}$.

The case of $\alpha = \infty$ (H_∞-optimization) is substantially more complicated and the corresponding optimization problem in Eq. 4-6 is refered to in this report as the "general distance problem" (GDP). This terminology arises from viewing the optimization in 4-6 as follows:

Given $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \in L_\infty$, find the optimal $Q \in H_\infty$ such that

$$\gamma_o = \min_{Q \in H_\infty} \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty \tag{4-11}$$

is minimized. Note that the minimum norm is the distance

$$\gamma_o = dist\left(R, \begin{bmatrix} H_\infty & 0 \\ 0 & 0 \end{bmatrix} \right) \tag{4-12}$$

from R to the set of (matrix) functions of the form

$$\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}, \quad Q \in H_\infty. \tag{4-13}$$

This class of problems will be called the "4-block problem" in this report to distinguish from the special case where $\begin{bmatrix} R_{21} & R_{22} \end{bmatrix}$ or $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}$ is identically zero. The latter will be referred as the "1-block problem". If both $\begin{bmatrix} R_{21} & R_{22} \end{bmatrix}$ and $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}$ are zero, this is known as the "best (or Hankel) approximation" problem (eg. Adamjan, Arov, and Krein (1971,1978), Sarason (1967), Glover (1984)) since

$$\min_{Q \in RH_\infty} \| R_{11} - Q \|_\infty = \| H_{R_{11}} \|. \tag{4-14}$$

This problem has the special property that it reduces to a finite dimensional eigenvalue problem and solutions can be computed with standard matrix algebra routines. Note that, in general, a Q_{opt} which achieves the minimum in 4-14 does not minimize the GDP in 4-11, although it will be shown to provide a reasonable approximation to the actual minimizer.

1.5 μ For Complex Perturbations

In the previous sections, it was shown that robust performance and stability with structured uncertainty reduces to computing μ for constant matrices $G(j\omega)$ and then taking sup over all ω. For this to be useful, we must have ways of computing μ or bounds for it. This section will begin by outlining some of the mathematical properties of μ for complex perturbations and viewing it as a natural generalization of the spectral radius ρ, and the spectral norm (maximum singular value) $\bar{\sigma}$. The rest of this section will focus on using scalings to characterize μ in terms of ρ and $\bar{\sigma}$.

Suppose that Δ is some subalgebra of matrices satisfying

$$\{\lambda I | \lambda \in C\} \quad \subset \quad \Delta \quad \subset \quad C^{N \times N}. \tag{5.1}$$

In this report we will be interested in block diagonal Δ. Define the spectrum, $sp(M)$, and inverse spectrum, $isp(M)$, of a matrix $M \in C^{N \times N}$ *with respect to the subalgebra* Δ as

$$sp(M) = \left\{ \Delta \in \Delta \;\middle|\; \det(M - \Delta) = 0 \right\}$$

$$isp(M) = \left\{ \Delta \in \Delta \;\middle|\; \det(I - M\Delta) = 0 \right\}. \tag{5.2}$$

Since both sets depend on Δ it would be appropriate to subscript the symbols, but to keep notation simple this will be avoided throughout. The set $sp(M)$ is a natural generalization of the usual notion of spectrum and is always nonempty. In this context, μ can be viewed as a natural generalization of spectral radius since it is easily verified that

$$\mu(M) = \sup_{\Delta \in sp(M)} \underline{\sigma}(\Delta). \tag{5.3}$$

If $\mu(M) \neq 0$ (which is equivalent to $isp(M) \neq \emptyset$) then

$$\mu(M) = \sup_{\Delta \in isp(M)} \frac{1}{\bar{\sigma}(\Delta)}. \tag{5.4}$$

This characterization emphasizes the view of μ as a generalization of $\bar{\sigma}$ and is simply a restatement of (3.4). Indeed, in the special cases where Δ is equal to one of its possible extreme sets in (5.1), μ is exactly either the usual spectral radius or maximum singular value:

$$\begin{aligned}
\Delta &= \{\lambda I | \lambda \in C\} \quad &\Rightarrow \quad \mu(M) &= \rho(M) \\
\Delta &= C^{N \times N} \quad &\Rightarrow \quad \mu(M) &= \bar{\sigma}(M)
\end{aligned} \tag{5.5}$$

It is possible to use these two special cases to obtain bounds for μ. For any set $\boldsymbol{\Delta}$ it easy to see that

$$\rho(M) \quad \leq \quad \mu(M) \quad \leq \quad \sigma(M) \tag{5.6}$$

but these bounds are not directly useful for computation as matrices may be found that make the differences between the bounds and μ as large as desired.

It is possible to improve the bounds in (5.6) by using simple properties of $\boldsymbol{\Delta}$. Suppose that $\mathcal{U}$ and $\mathcal{D}$ are sets such that for any $\Delta \in \boldsymbol{\Delta}$

$$U \in \mathcal{U} \quad \Rightarrow \quad \sigma(U\Delta) = \sigma(\Delta)$$

$$D \in \mathcal{D} \quad \Rightarrow \quad D^{-1}\Delta D = \Delta. \tag{5.7}$$

Then it is easy to see from the definition of μ that

$$U \in \mathcal{U} \quad \Rightarrow \quad \mu(MU) = \mu(M)$$

$$D \in \mathcal{D} \quad \Rightarrow \quad \mu\left(DMD^{-1}\right) = \mu(M) \tag{5.8}$$

so the bounds in (5.6) can be improved to

$$\sup_{U \in \mathcal{U}} \rho(MU) \leq \mu(M) \leq \inf_{D \in \mathcal{D}} \sigma\left(DMD^{-1}\right). \tag{5.9}$$

The key theorems about μ show when these inequalities are actually equalities.

Let us first consider the case where all the blocks are complex and none are repeated. Then we have the sets

$$\boldsymbol{\Delta} = \left\{ \operatorname{diag}\left(\Delta_1, \Delta_2, \ldots, \Delta_n\right) \,\middle|\, \Delta_j \in \mathbf{C}^{m_j \times m_j} \right\}$$

$$\mathcal{U} = \left\{ \operatorname{diag}\left(U_1, U_2, \ldots, U_n\right) \,\middle|\, U_j^* U_j = I \right\}$$

$$\mathcal{D} = \left\{ \operatorname{diag}\left(d_1 I, d_2 I, \ldots, d_n I\right) \,\middle|\, d_i \in \mathbf{R}_+ \right\} \tag{5.10}$$

It is easy to verify that (5.7) holds so that the inequalities in (5.9) apply. What is more important is that

$$\sup_{U \in \mathcal{U}} \rho(MU) = \mu(M) \tag{5.11}$$

holds for all M and $\boldsymbol{\Delta}$ and

$$\mu(M) = \inf_{D \in \mathcal{D}} \sigma\left(DMD^{-1}\right) \tag{5.12}$$

if $n \leq 3$ (three or fewer blocks) [D13]. There are other conditions under which this upper bound is an equality but they are more cumbersome to state and generally of less interest. An example of strict inequality for the upper bound has been found for $n = 4$. Extensive computational experimentation has yet to find a matrix for which the upper bound exceeds μ (actually some lower bound for μ) by more than 15%, and the upper bound is nearly equal for most matrices. This seems to be independent of matrix size and number of blocks. This is encouraging but additional theoretical work is needed to guarantee the quality of the upper bound in general.

The case of repeated blocks is less well understood. To see what U and D arise when there are repeated blocks consider the simple case where each block is a repeated scalar

$$\Delta = \left\{ \text{diag} \left(\delta_1 I, \delta_2 I, \ldots, \delta_n I \right) \,\middle|\, \delta_j \in C \right\}$$

$$U = \left\{ \text{diag} \left(u_1 I, u_2 I, \ldots, u_n I \right) \,\middle|\, u_j \in C, |u_j| = 1 \right\}$$

$$D = \left\{ \text{diag} \left(D_1, D_2, \ldots, D_n \right) \,\middle|\, D_j \quad \text{invertible} \right\} \tag{5.13}$$

It is possible to restrict the $D \in D$ to positive definite Hermitian matrices $(D_j = D_j^* > 0)$ without loss of generality. As above, the inequalities in (5.9) hold and (5.11) also holds for all M and Δ. Unfortunately, the conditions under which the upper bound is an equality are not easily checked. The computational experience with the case of repeated blocks is much more limited than with nonrepeated blocks, but the evidence so far suggests that the upper bound is also nearly an equality. The case of repeated nonscalar blocks is just the obvious combination of the the above two cases.

The lower bounds in terms of $\rho(MU)$ have the desirable property of always achieving μ independent of the number of blocks. Unfortunately, $\rho(MU)$ can have multiple local maxima which are not global so direct computation of (5.11) by gradient search may not find the actual maximum. At this time there is no alternative scheme guaranteed to find the global maximum that has reasonable computational properties. Fan and Tits (1985) do have an alternative scheme for a lower bound which does not guarantee that μ will be found but appears to be very fast and has many advantages over using (5.11).

The upper bound in (5.9) is more easily found since the expression $\bar{\sigma} \left(DMD^{-1} \right)$ has only global minima. This is a direct consequence of the fact that $\bar{\sigma} \left(e^D M e^{-D} \right)$ is convex in D. This fact was used in [D13] to

argue that the upper bound in (5.9) probably offered a reasonable alternative to (5.11) for computation of μ. The original proof of convexity was rather cumbersome and appeared later in [S5]. A much simpler proof is included at the end of this section.

Computational experience to date has indicated that it is desirable in practice to use both upper and lower bounds for μ, since the existing bounds nicely complement each other. The upper bound is easily computed but may not give μ except in special cases. On the other hand, it appears to be nearly equal to μ in all cases. The existing lower bounds (including both (5.11) and those Fan and Tits) are, in principle, equal to μ in all cases but may fail because of local maxima. By having an upper bound it is much easier to recognize when a local maxima is not global and restart the algorithm with another initial guess. Extensive computational experience has yet to reveal a (complex) μ problem where the bounds obtained in this way differed by more than about 15%. More research is needed to show whether this is always true. It could simply turn out that counterexamples exist but are difficult to find.

Proof of convexity of $\bar{\sigma}\left(e^D M e^{-D}\right)$

This proof is based on the following simple lemma. Suppose that $f : \mathbf{R} \mapsto \mathbf{R}$ is continuous.

Lemma. Suppose $\forall x, \exists g_x : \mathbf{R} \mapsto \mathbf{R}$ twice differentiable such that $f(x) = g_x(x)$, $f(t) \geq g_x(t)$ and $\frac{d^2}{dt^2} g_x(t)\Big|_{t=x} \geq 0$. Then f is convex.

To prove convexity of $\bar{\sigma}\left(e^D M e^{-D}\right)$ it suffices to prove convexity of $f(x) = \bar{\sigma}\left(e^{Dx} M e^{-Dx}\right)$ for arbitrary $D = D^* \in \mathbf{C}^{N \times N}$. To apply the lemma define $M_x = e^{Dx} M e^{-Dx}$ and let u and v be (any) singular vectors such that $f(x) = \bar{\sigma}(M_x) = u^* M_x v$. Then define $g_x(t) = \Re\left\{u^* e^{Dt} M e^{-Dt} v\right\}$. Since $f(t) \geq g_x(t)$ and

$$
\begin{aligned}
\frac{d^2}{dt^2} g_x(t)\Big|_{t=x} &= \Re\left\{u^* \left(D^2 M_x - 2D M_x D + M_x D\right) v\right\} \\
&= f(x)(u^* D^2 u + v^* D^2 v) - 2\Re\left\{u^* D M_x D v\right\} \\
&= [u^* D \quad v^* D] \begin{bmatrix} f(x)I & -M_x \\ -M_x & f(x)I \end{bmatrix} \begin{bmatrix} Du \\ Dv \end{bmatrix} \\
&\geq 0
\end{aligned}
$$

by the lemma f is convex.

1.6 Computation of μ for Real Perturbations

The properties of μ when Δ has some elements restricted to be real are quite different from the purely complex case. Suppose that

$$\Delta = \{\text{diag}\,(\delta_1, \delta_2, \ldots, \delta_m, \Delta_1, \Delta_2, \ldots, \Delta_n)$$

$$\left| \; \delta_i \in \mathbf{R}, \Delta_j \in C^{k_j \times k_j} \right\} \tag{6.1}$$

and $sp(M)$ and $isp(M)$ are defined for Δ exactly as in (5.2). In this case, it is possible for either $sp(M)$ or $isp(M)$ to be the empty set. Furthermore, (5.3) is no longer a correct characterization of μ in general and there is no natural way to view μ as a simple generalization of the usual notion of spectral radius. Of course, (5.4) still applies provided $\mu(M) \neq 0$ (i.e. $isp(M) \neq \emptyset$).

This section will focus on upper bounds to μ that can be obtained by scaling σ. The choice of scaling is based on the following lemma which characterizes a useful class of scalings. In the following lemma, assume that

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

and $\det(I - T_{22}M) \neq 0$ so that

$$F_l(T, M) = T_{11} + T_{12}\,M(I - T_{22}M)^{-1}\,T_{21}$$

is well-defined.

Lemma:

$$\text{Suppose } \exists T \text{ such that } \mathbf{B}\Delta \subset \left\{ F_u(T, \Delta) \;\middle|\; \sigma(\Delta) < 1 \right\}$$

$$\text{then } \sigma(F_l(T, M)) \leq 1 \Rightarrow \mu(M) \leq 1$$

This lemma says that if $F_u(T, \Delta)$ "covers" $\mathbf{B}\Delta$ then T can be used to obtain an upper bound for μ. The proof is given at the end of this section. The next step is to identify a set of T's that satisfy the lemma. To this end define

$$D = \{\text{diag}\,(d_1, d_2, \ldots, d_n, d_{m+1}I, d_{m+2}I, \ldots, d_{m+n}I)$$

$$\left|\, d_i \,\in\, \mathbf{R_+} \right\}$$

$$\mathcal{C} = \left\{ \text{diag}\,(c_1, c_2, \ldots, c_m, 0, 0, \ldots, 0) \,\middle|\, c_i \,\in\, [-1, 1] \right\}$$

$$\mathcal{T} = \left\{ \left[\begin{array}{cc} jC & (I - C^2)^{1/2} D \\ D^{-1} & 0 \end{array} \right] \,\middle|\, D \in \mathcal{D}\,,\, C \in \mathcal{C} \right\} \tag{6.2}$$

where $\mathcal{D}$ and $\mathcal{C}$ are partitioned conformally with $\mathbf{\Delta}$. With these definitions, for $T \in \mathcal{T}$

$$F_l(T, M) = jC + (I - C^2)^{1/2}\, DMD^{-1}. \tag{6.3}$$

It is a matter of some simple algebra to show that all $T \in \mathcal{T}$ satisfy the lemma. Note that if there are no real parameters ($m=0$), then $F_l(T, M) = DMD^{-1}$ and this scaling reduces to that considered in (5.10) and (5.12). Other $\mathcal{T}$ also satisfy the lemma but this parametrization is convenient because (6.3) is particularly simple.

Obtaining an upper bound based on (6.3) is somewhat more complicated than is possible in (5.12). The difficulty is that the above lemma only implies that $\mu(M) \leq 1$ and does not scale. Using the sets in (6.2), we can define

$$\tilde{\mu}(M) = \inf_{\alpha \in \mathbf{R_+}} \left\{ \alpha \,\middle|\, \inf_{T \in \mathcal{T}} \bar{\sigma} \left(F_l(T, \frac{1}{\alpha} M) \right) < 1 \right\}. \tag{6.4}$$

It follows immediately from the lemma that

$$\mu(M) \ \leq \ \tilde{\mu}(M)$$

and thus $\tilde{\mu}$ provides an upper bound for μ. Again note that for no real parameters, $\tilde{\mu}(M)$ simplifies to

$$\tilde{\mu}(M) = \inf_{D \in \mathcal{D}} \bar{\sigma} \left(DMD^{-1} \right)$$

The natural question is how good a bound is $\tilde{\mu}$ for μ. Recall that for $m = 0$, $n \leq 3$ that $\tilde{\mu} = \mu$ for all matrices independent of block size (m is the number of real parameters and n is the number of complex blocks). A simple extension of this result yields $\tilde{\mu} = \mu$ when $m = 1$, $n \leq 2$. Although counterexamples exist for problems with more than these number of blocks, experience has shown that $\tilde{\mu}$ is often a good approximation to μ even in these cases. While this experience is encouraging it is not conclusive and additional research is needed to establish the value of $\tilde{\mu}$. Unfortunately, when there is more than one real parameter it is possible for $\mu(M) \ll \tilde{\mu}(M)$.

Proof of lemma: Since

$$\det(I - T_{22}M)\det\left(I - F_l(T,M)\Delta\right)$$

$$= \det(I - T_{11}\Delta)\det\left(I - MF_u(T,\Delta)\right),$$

$$\sigma\left(F_l(T,M)\right) \leq 1 \Rightarrow \det(I - F_l(T,M)\Delta) \neq 0 \ \forall \sigma(\Delta) < 1$$

$$\Rightarrow \det(I - M\ F_u(T,\Delta)) \neq 0 \ \forall \sigma(\Delta) < 1$$

$$\Rightarrow \det(I - M\Delta) \neq 0 \ \forall \Delta \in B\boldsymbol{\Delta}$$

$$\Rightarrow \mu(M) \leq 1$$

1.7 μ-Synthesis

The previous sections on analysis showed that the synthesis problem reduces to finding a stabilizing controller K so that

$$\|F_l(P,K)\|_\alpha \leq 1 \qquad \alpha = \infty \text{ or } \mu \tag{7.1}$$

where $F_l(P,K) = P_{11} + P_{12} K(I - P_{22}K)^{-1} P_{21}$. The solution to (7.1) for $\alpha = \infty$ was outlined in section 4 and additional details are presented in the remaining chapters of this report. This H_∞-synthesis solution can be used to provide an approach to solving the μ-norm synthesis problem, refered to as μ-synthesis.

Recall that the first step in the H_∞ synthesis solution involves finding J so that the substitution $K = F_l(J,Q)$ yields

$$F_l(P,K) = F_l(P,F_l(J,Q)) = R + UQV \tag{7.2}$$

with $F_l(P,K)$ internally stable iff $Q \in H_\infty$. Further, U is inner and V co-inner ($U^*U = I$ and $VV^* = I$), and there exist complementary inner factors $U_\perp$ and $V_\perp$ such that $[UU_\perp]$ and $\begin{bmatrix} V \\ V_\perp \end{bmatrix}$ are both square and inner. The U and V are obtained from coprime factorizations $P_{12} = UM_1^{-1}$ and $P_{21} = M_2^{-1}V$. The next step involves using a rational matrix version of the Davis-Kahan-Weinberger matrix dilation results [D1] to further reduce the problem to one of finding $\hat{Q} \in RH_\infty$ such that

$$\|G + \hat{Q}\|_\infty \leq 1 \tag{7.3}$$

where $G \in RL_\infty$. This problem can then be solved using the Hankel norm approximation methods developed by Glover [G2]. The resulting optimal $\hat{Q}$ can then be used to find first the optimal Q and then the optimal K.

The μ-synthesis problem does not yet have as complete a solution as does the H_∞ synthesis problem. A reasonable approach would be to try to find a stabilizing controller K and scaling D so that

$$\|DF_l(P,K)D^{-1}\|_\infty \leq 1. \tag{7.4}$$

One method to do this is to alternately minimize the above expression for either K and D while holding the other constant. For fixed D the left-hand side of (7.4) is just an H_∞ control problem and can be solved using the methods reviewed above. For fixed K, the left-hand side of (7.4) can be minimized at each frequency

as a convex optimization problem in D. The resulting D can be fit with a stable, rational transfer function with stable inverse (the phase of D does not affect the norm).

This approach to μ-synthesis has been successfully applied to several example problems. In principle, it could be used to obtain controllers that are arbitrarily close to μ-optimal in the case of 3 or fewer blocks and provide nearly optimal controllers for the general case. This would depend on the suggested iterative scheme converging to the global optimal K and D. Unfortunately, individual convexity in the two parameters of an optimization problem does not imply joint convexity, and this scheme is not always guaranteed to converge globally to the best K and D.

To better understand the properties of the problem in (7.4) it is useful to consider the constant matrix problem. Using (7.2), we can reduce (7.2) to

$$\|D(R + UQV)D^{-1}\|_\infty \leq 1. \tag{7.5}$$

for constant R, U, V with $U^*U = 1$ and $VV^* = 1$. For $D = I$, it follows from [G2] that

$$\min_Q \bar{\sigma}(R + UQV) = \max\left(\bar{\sigma}(U_\perp^* R), \bar{\sigma}(RV_\perp^*)\right) \tag{7.6}$$

where $U_\perp^*$ and $V_\perp^*$ are chosen so that $[U \; U_\perp]$ and $\begin{bmatrix} V \\ V_\perp \end{bmatrix}$ are both square and unitary. All of these quantities are easily computed using standard SVD routines.

Posing (7.5) as an optimization problems gives

$$\min_{D,Q} \bar{\sigma}\left(D(R + UQV)D^{-1}\right). \tag{7.7}$$

It is known that this problem is convex in either D (actually $\ln(D)$) or Q individually when the other is held fixed, but is not convex in both variables jointly. This means that the iterative scheme suggested as a possible approach to μ-synthesis is not guaranteed to converge even in the constant matrix case. It is possible, however, to compute the desired D in (7.7) directly.

The result in (7.6) may be applied to (7.8) to obtain

$$\min_Q \bar{\sigma}\left(D(R + UQV)D^{-1}\right) =$$

$$\max\left\{\bar{\sigma}\left((DU)_\perp^* DRD^{-1}\right), \bar{\sigma}\left(DRD^{-1}(VD^{-1})_\perp^*\right)\right\} \tag{7.8}$$

where

$$(DU)_\perp \triangleq D^{-1}U_\perp (U_\perp^* D^{-2} U_\perp)^{-1/2} \tag{7.9}$$

and $(VD^{-1})_\perp$ is defined similarly. Note that $\left[DU(U^*D^2U)^{-1/2} \;\; (DU)_\perp\right]$ is unitary. It can be shown that the right hand side of (7.8) is convex in $\ln(D)$ so that the "optimal" scaling for (7.7) may be computed by search in advance. This gives a tight lower bound for (7.7) and the resulting D scaling may be used to compute the optimal Q.

A simple example will illustrate all the essential features of this possibly confusing sequence of ideas. Consider the problem

$$\min_q \; \mu \left(\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} q[1 \; 0] \right)$$

$$= \min_{q,d} \; \bar\sigma \left(\begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} \right). \tag{7.10}$$

The μ-optimal q is $q = 0$ which gives $\mu = 1$. For fixed d the $\bar\sigma$-optimal $q = d^2$ and for fixed $q > 0$ the $\bar\sigma$-optimal d is $d = \sqrt{q}$. Thus, iteratively solving for either q or d will immediately converge to the curve $q = d^2$. For example, with the initial guess of $q = d = 1$, the iterative scheme will not change either q or d and will thus fail to find the global optimum.

On the other hand,

$$\min_q \; \bar\sigma \begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} = \max \left(\bar\sigma([-1 \; d]) \, , \, \bar\sigma \left(\begin{bmatrix} d \\ 1 \end{bmatrix} \right) \right)$$

$$= \sqrt{1 + d^2}. \tag{7.11}$$

Thus,

$$\min_d \left(\min_q \; \bar\sigma \begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} \right) = \min_d \; \sqrt{1 + d^2} \tag{7.12}$$

which is clearly convex in both d and lnd and achieves its minimum as $d \to 0$. If the expression in (7.8) were used to compute the d in advance, it would be possible to find the optimal achievable level for (7.9). This example also illustrates why, strictly speaking, inf, not min must be used for the D scalings as in (5.9). This issue will not be taken up in this report. It turns out to be of little significance anyway.

The simplest application of these ideas to the selection of the D scalings for the μ-synthesis problem is to compute an initial guess for D at each frequency using (7.8). This would be the optimal D for an acausal

controller, and should provide a good initial guess for the optimal D for the causal controller problem. A deeper question is whether some generalization of (7.8) and its convexity properties applies to the rational case. While this seems likely, the details have not been worked out and the practical implications are uncertain. For some additional results on μ-synthesis, see [S5].

1.8 Outline of Chapters 2-8

The remainder of this report considers various issues arising in H_∞ optimal control theory which are associated with a particular solution approach which involves reduction of the standard problem to a "general distance problem" (GDP) in L_∞. Since H_∞ methods were introduced to the engineering community by Helton (1981), Tannenbaum (1980), and Zames(1981), there have been numerous papers on the subject from many points of view (e.g. Chang and Pearson (1984), Feintuch and Francis (1984), Foo and Posthlethwaite (1984), Francis and Zames (1984), Glover (1984), Khargonekar and Tannenbaum (1985), Kwakernaak (1983), Safonov and Verma (1983), Verma and Jonkheere (1984), to name just a few). This paper will focus on the approach developed in Doyle (1983, 1984), Chu and Doyle (1984, 1985), and Chu (1985), which was outlined in Section 1.4. For an overview of H_∞ control theory and a review of the literature see the expository paper by Francis and Doyle (1985).

This report is divided into eight chapters. Chapter 2 expands on the overview in Section 1.4 of the general optimal synthesis theory from Doyle (1984), which includes both the H_2 and H_∞ optimal control problems. The affine parametrization of the closed-loop transfer matrix is obtained following Youla's parametrization of all controllers achieving internal stability. A particular parametrization is employed involving coprime factorization with inner numerator; the H_2-optimal controller is found immediately and the H_∞-optimization formulation is transformed to an equivalent "general distance problem" (GDP). Two simple examples are also presented.

Chapter 3 describes results on optimal solutions to the general distance problem. The optimal norm of the GDP can be expressed in terms of an induced operator norm or an equivalent eigenvalue problem involving a combination of *Hankel* and *Toeplitz* operators. The approach is conceptually elegant; however, it does not yield a computable formula for either the minimal norm or the optimal solution.

In Section 4, the approach of γ-iteration is introduced. It essentially involves guessing a γ and then reducing the problem to an equivalent best approximation formulation. The guess for γ is iterated until it converges to the minimal norm, and the optimal solution is thus obtained. Some fairly tight bounds for the minimal norm which are easily computed are also given; these immediately allow for reasonable estimates of the minimal norm as well as giving an approximation technique for obtaining suboptimal solu-

tions that are within a guaranteed bound of optimal. To study the convergence properties of the γ-iteration it is then viewed as a problem of finding the zero crossing of a function. It is established that this function is continuous, monotonically decreasing, convex, and in turn bounded by some very simple functions. These properties make it possible to obtain very rapid convergence of the γ-iteration. An interesting example is given to illustrate some important aspects of the general distance problem which were not previously well-understood.

The state-space formulation of the γ-iteration is then developed and presented in Chapter 5. In the γ-iteration, a key step is to find the spectral factor of a para-Hermitian matrix of the form $(\gamma^2 I - G^*G)$ (or $(\gamma^2 I - GG^*)$) which typically, requires one to solve for a coprime factorization with inner denominator and a standard spectral factorization. Each of the factorizations requires finding the stabilizing solution of an algebraic Riccati equation (ARE). The ARE associated with coprime factorization has a special structure where the constant term is identically zero. A very efficient algorithm based on a Schur decomposition is developed to solve the ARE with this special structure. For completeness, balanced realizations and Glover's algorithm to the best approximation problem are also reviewed. Combining previous results, the "closed-form" state-space optimal solution of the general distance problem is then obtained.

Chapter 6 recapitulates the results detailed in Chapter 2 through Chapter 5, with a discussion of some numerical aspects of the algorithms. Chapter 6 also discusses some results on model reduction in the context of the **GDP** as a method for obtaining suboptimal solutions with reduced order. The focus here is on model reduction techniques where the error produced by a reduction can be related directly to the degree of suboptimality of the resulting solution.

CHAPTER 2

OPTIMAL SYNTHESIS THEORY

This chapter contains a review of the general H_2 and H_∞ optimal synthesis theory. It is an expansion of the material outlined in Section 4 of Chapter 1 and shows in detail how the general H_∞ problem reduces to the "General Distance Problem". Recall that the framework used here is depicted in Figure 2-1. The approach used in this chapter follows closely that in [D16].
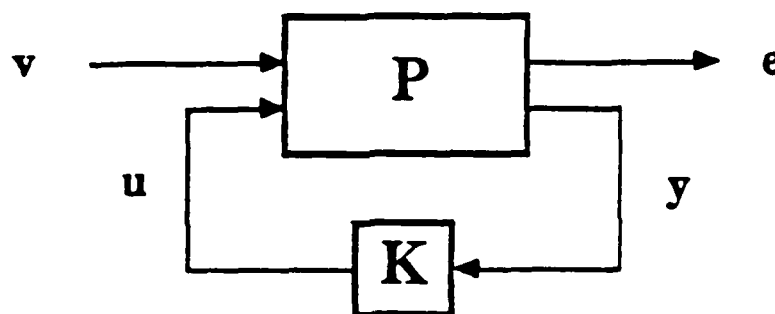
Figure 2-1. General Framework for Synthesis

Partitioning P appropriately, the closed-loop transfer function matrix from v to e can be writtened as the following LFT

$$\mathbf{F}_l(P,K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21} \qquad (2\text{-}1)$$

where

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in R_p^{(p_1 + p_2) \times (m_1 + m_2)} \quad , \quad P_{ij} \in R^{p_i \times m_j} . \tag{2-2}$$

The problem in H_2 and H_∞ synthesis is to find a controller K achieving "internal stability" such that the performance measure $\| \mathbf{F}_l(P,K) \|_\alpha$ is minimized for $\alpha = 2$ or ∞.

In this chapter, the notion of internal stability for the LFT $\mathbf{F}_l(P,K)$ is reviewed first in Section 2.1. Section 2.2 gives an algebraic treatment of the Youla parametrization of all stabilizing controllers in terms of a stable parameter matrix $Q \in RH_\infty$ with the LFT representation,

$$K = \mathbf{F}_l(J,Q) . \tag{2-3}$$

This parametrization has the additional property that substitution of Eq. 2-3 into Eq. 2-1 yields

$$\mathbf{F}_l(P,K) = \mathbf{F}_l(P,\mathbf{F}_l(J,Q))$$
$$= \mathbf{F}_l(T,Q)$$

where

$$\mathbf{F}_l(T,Q) = T_{11} + T_{12}QT_{21} \tag{2-4}$$

providing an affine parametrization of all internally stable closed-loop transfer function matrices $\mathbf{F}_l(P,K)$. The Youla parameterization is then constructed using the standard state-space computations of observer-based stabilization methods, providing explicit realizations of the desired J in Eq. 2-3 in terms of a realization of P. The affine parameterization of the closed-loop system is also derived using a state-space realization.

In Section 2.3, a particular parametrization is derived such that T_{12} is inner and T_{21} is co-inner. The algebraic Riccati equation plays an essential role in obtaining such a parametrization. Using the unitary-invariant property of the H_2 and H_∞-norm, the H_2-optimal controller is immediately obtained. The H_∞ optimal control formulation is transformed into an equivalent "general distance problem" which will be discussed in great detail in the next three chapters. Section 2.4 presents two examples to illustrate the H_2 and H_∞ optimal control respectively.

2.1 Internal Stability

The results in this section are entirely standard, although typically not expressed in the LFT framework. For an interesting alternative treatment which is closely related see Nett (1985). In this section P and K are fixed proper transfer function matrices. The block diagram associated with Figure 2-1 represents the two equations

$$
\begin{bmatrix} e \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} , \quad u = Ky .
$$

It is convenient to introduce two fictitious external signals, w_1 and w_2, as in Figure 2-2.
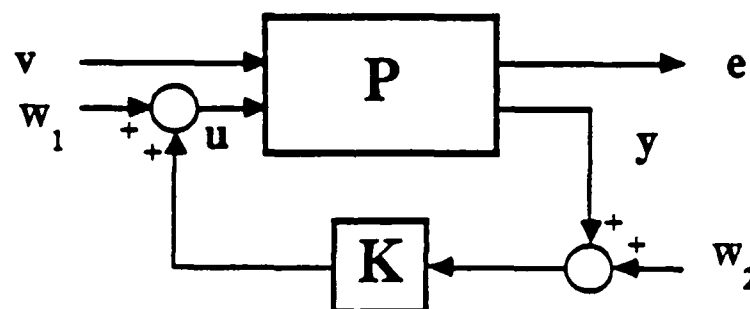
Figure 2-2.

Suppose the signals v, w_1, and w_2 are specified and that u in Figure 2-2 is well-defined. Then so are e and y. Thus it makes sense to define the system as diagrammed in Figure 2-2 to be "well-posed" provided the transfer function matrix from $\begin{bmatrix} v^T & w_1^T & w_2^T \end{bmatrix}^T$ to u exists and is a proper one. The following theorem shows the necessary and sufficient conditions of well-posedness.

Theorem 1

The following statements are equivalent:

(i) The system as diagrammed in Figure 2-1 is well-posed.

(ii) $I - K(\infty)P_{22}(\infty)$ is invertible. (2-5)

(iii) $\begin{bmatrix} I & -K(\infty) \\ -P_{22}(\infty) & I \end{bmatrix}$ is invertible. (2-6)

(iv) $I - P_{22}(\infty)K(\infty)$ is invertible. (2-7)

Alternatively, the well-posedness condition can be stated in terms of state-space realizations. For this purpose, introduce minimal realizations of P and K:

$$P = \begin{bmatrix} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix}$$ (2-8)

$$K = \begin{bmatrix} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{bmatrix}.$$ (2-9)

The partition in Eq. 2-8 corresponds to that in Eq. 2-2, i.e.,

$$P_{ij} = \begin{bmatrix} A & B_j \\ \hline C_i & D_{ij} \end{bmatrix}.$$ (2-10)

Then $P_{22}(\infty) = D_{22}$ and $K(\infty) = \hat{D}$, and so (from Eq. 2-6) well-posedness is equivalent to the condition that the matrix

$$\begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}$$ (2-11)

is invertible. Well-posedness of the system as represented by its transfer function matrix will be assumed throughout this chapter.

Let z and $\dot{z}$ denote the state vectors associated with minimal realizations of P and K respectively, and write the corresponding system equations for the interconnection structure in Figure 2-1 with v set to zero and e ignored as:

$$\dot{z} = Az + B_2 u \tag{2-12a}$$

$$y = C_2 z + D_{22} u \tag{2-12b}$$

$$\dot{\hat{z}} = \hat{A} \hat{z} + \hat{B} y \tag{2-12c}$$

$$u = \hat{C} \hat{z} + \hat{D} y. \tag{2-12d}$$

The system diagrammed in Figure 2-1 is "internally stable" provided the null solution $(z, \hat{z}) = (0,0)$ of Eqs. 2-12a to 2-12d is asymptotically stable To get a concrete characterization of internal stability, solve Eqs. 2-12b and 2-12d for u and y:

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} z \\ \hat{z} \end{bmatrix}.$$

(Note that the inverse exists because of the well-posedness condition Eq. 2-11). Now substitute this expression for u and y into Eqs. 2-12a and 2-12c to get

$$\frac{d}{dt} \begin{bmatrix} z \\ \hat{z} \end{bmatrix} = \bar{A} \begin{bmatrix} z \\ \hat{z} \end{bmatrix}$$

where

$$\bar{A} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C_2 & 0 \end{bmatrix}.$$

Thus internal stability is equivalent to the condition that $\bar{A}$ is a stability matrix, i.e., all eigenvalues of $\bar{A}$ lie in the open *lhp* .

It is not difficult to verify that the above definition of internal stability depends only on the transfer function matrices P and K, and not on the specific minimal realizations of them. The following result is standard.

Theorem 2

Consider a minimal realization of the system P as in Eq. 2-8. Then there exists a proper real-rational transfer function K achieving internal stability if and only if the pair (A, B_2) is stabilizable and the pair (C_2, A) is detectable. The latter stabilizability and detectability conditions are assumed throughout this chapter.

Since

$$P_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right] , \qquad (2\text{-}13)$$

Eqs. 2-12a through 2-12d constitute a state-space representation of the system diagrammed in Figure 2-3. Although the realization in Eq. 2-13 is not necessarily minimal, it is stabilizable and detectable, and these are enough to yield the following result.
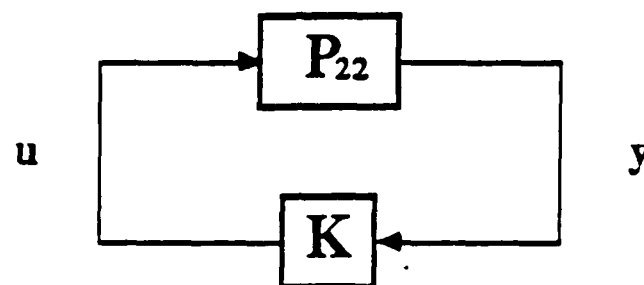
Figure 2-3.

Theorem 3

The system diagrammed in Figure 2-1 is internally stable if and only if the system diagrammed in Figure 2-3 is internally stable.

The above notion of internal stability was defined in terms of state-space realizations of P_{22} and K. It is also important and useful to characterize internal stability from an input/output point of view. For this, consider the feedback system diagrammed in Figure 2-4. This system has an input/output relationship:

$$\begin{bmatrix} I & -K \\ -P_{22} & I \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \tag{2-14}$$
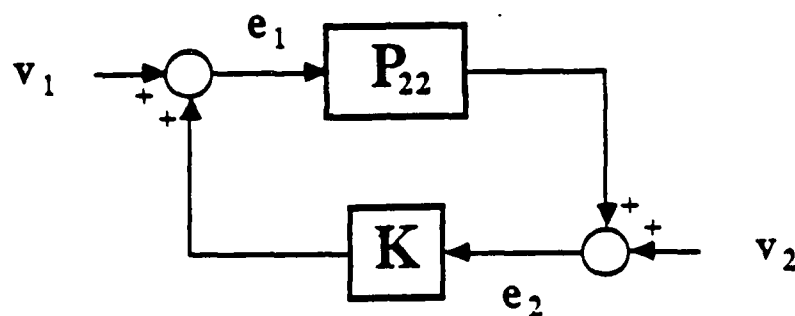
Figure 2-4.

It is intuitively clear that if the system diagrammed in Figure 2-4 is internally stable, then for all bounded inputs (v_1, v_2), the outputs (e_1, e_2) are also bounded. This idea leads to an input/output characterization of internal stability.

Theorem 4

The system diagrammed in Figure 2-4 is internally stable if and only if $(I - P_{22}K)$ is invertible and the transfer function matrix

$$\begin{bmatrix} I & -K \\ -P_{22} & I \end{bmatrix}^{-1} = \begin{bmatrix} I + K(I - P_{22}K)^{-1}P_{22} & K(I - P_{22}K)^{-1} \\ (I - P_{22}K)^{-1}P_{22} & (I - P_{22}K)^{-1} \end{bmatrix} \tag{2-15}$$

between (v_1, v_2) and (e_1, e_2) belongs to RH_∞.

Note that to check internal stability it is necessary (and sufficient) to check that each of the four transfer function matrices in Eq. 2-15 are in RH_∞. It is not difficult to construct examples involving P_{22} and K such that some of the four transfer matrices in Eq. 2-15 are in RH_∞ while the others are unstable [D4].

2.2 Parametrization of All Stabilizing Controllers

For the discussion here, there are two main approaches to constructing stabilizing controllers for linear systems: the Youla parametrization and state-space methods using observers and state feedback [K2,L4]. Each is well-known among the control community and each has its advantages. The Youla parametrization yields *all* stabilizing controllers as well as a convenient affine parametrization of the closed-loop system. Unfortunately, the standard algebraic treatment of this subject gives no reliable scheme to compute the coefficients of the parametrization. Observer-based stabilizing controllers, on the other hand, are easily constructed in terms of a realization of the transfer function matrix P using a variety of state-space computation schemes.

In has been shown that these two methods of stabilization are actually equivalent. This allows the Youla parametrization to be constructed using the standard state-space computations of observer-based stabilization methods, providing explicit realizations of the desired J in Eq. 2-3 in terms of a realization of P. Combining the results of the previous section and this section, the desired affine parametrization of the closed-loop system is then obtained.

It should be noted that many of the results on the connections between the algebraic and observer-based stabilization methods were discovered independently by Nett and coauthors [N2]. Also, many of these results were known within the "systems over rings" community [K4]. A complete treatment on the equivalence of these two stabilization methods and the parametrization in terms of the general framework (Figure 2-1) using linear fractional transformation was first given by Doyle [D16].

The following definitions of coprimeness in RH_∞ provide the appropriate framework for the equivalences to be discussed.

Definition: (Right Coprimeness)

Two matrices N, $M \in RH_\infty$ with the same number of columns m are right coprime if there exists X, $Y \in RH_\infty$ such that

$$XM + YN = I_m .$$

<div align="right">(2-16)</div>

Definition: (Left Coprimeness)

Two matrices $\tilde{N}$, $\tilde{M} \in RH_\infty$ with the same number of rows p are left coprime if there exists $\tilde{X}$, $\tilde{Y} \in RH_\infty$ such that

$$\tilde{M}\tilde{X} + \tilde{N}\tilde{Y} = I_p . \tag{2-17}$$

Remarks

(1) Eq. 2-16 (2-17) is equivalent to saying that the combined matrix $\begin{bmatrix} M \\ N \end{bmatrix}$ ($\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$) has a left (right) inverse in RH_∞.

(2) An alternative definition for the right (left) coprime factorization, rcf (lcf), is that two matrices in RH_∞ are right (left) coprime if every common right (left) divisor in RH_∞ is invertible in RH_∞. This is equivalent to the above definition in terms of a left inverse.

(3) Eq. 2-16 (or Eq. 2-17) is often called a Bezout (or Diophantine) identity.

It is a fact that every $G \in R_p$ (proper, real-rational transfer function matrices) has an rcf $G = NM^{-1}$ where the pair N, $M \in RH_\infty$ are right coprime. Similarly, there exists an lcf , defined in the obvious way by duality. The proof of the existence of such coprime factorizations can be found in several publications [D16,K4,N2,V3] with explicit realizations for the factorizations. In this section, it will be shown how these factorizations can be used to obtain a parametrization of all stabilizing controllers.

Beginning with rcf s and lcf s of P_{22} and K in Figure 2-4:

$$P_{22} = NM^{-1} = \tilde{M}^{-1}\tilde{N} , \tag{2-18}$$

$$K = UV^{-1} = \tilde{V}^{-1}\tilde{U} . \tag{2-19}$$

The following lemma is well-known.

Lemma 1

Consider the feedback system diagrammed as in Figure 2-4. The following conditions are equivalent:

(i)   The feedback system as represented by Eqs. 2-12a to 2-12d is internally stable.

(ii) $\begin{bmatrix} M & U \\ N & V \end{bmatrix}$ is invertible in RH_∞.

(iii) $\begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{bmatrix}$ is invertible in RH_∞.

(iv) $\tilde{V}M - \tilde{U}N$ is invertible in RH_∞.

(v) $\tilde{M}V - \tilde{N}U$ is invertible in RH_∞.

Explicit realizations for N, M, $\tilde{N}$, $\tilde{M}$, $U_\bullet$, $V_\bullet$, $\tilde{U}_\bullet$, and $\tilde{V}_\bullet$ will be given later in this section which satisfy Eq. 2-18 and

$$\begin{bmatrix} \tilde{V}_\bullet & -\tilde{U}_\bullet \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U_\bullet \\ N & V_\bullet \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{2-20}$$

which is often also called "doubly Bezout identity". By the above lemma

$$K_\bullet \triangleq U_\bullet V_\bullet^{-1} = \tilde{V}_\bullet^{-1}\tilde{U}_\bullet \tag{2-21}$$

then qualifies as a particular controller achieving internal stability. The result of the next theorem means that all stabilizing controllers can be expressed in terms of matrix $K_\bullet$ and a parameter matrix $Q \in RH_\infty$. The proof can be found elsewhere, for example, [D5,D16,V2,V3].

Theorem 2

The set of all proper controllers achieving internal stability for the feedback system (see Figure 2-1) is parametrized by the formula

$$K = (U_\bullet + MQ)(V_\bullet + NQ)^{-1} \tag{2-22a}$$

$$= (\tilde{V}_\bullet + Q\tilde{N})^{-1}(\tilde{U}_\bullet + Q\tilde{M}) \tag{2-22b}$$

$$= K_{\bullet} + \tilde{V}_{\bullet}^{-1} Q (I + V_{\bullet}^{-1} N Q)^{-1} V_{\bullet}^{-1} \qquad \text{(2-22c)}$$

where Q ranges over RH_∞ such that $(I + V_{\bullet}^{-1} N Q)(\infty)$ is invertible.

It is not difficult to recognize that K, as in Eq. 2-22c, can be expressed in terms of a LFT as shown in Figure 2-5, i.e.,

$$K = \mathbf{F}_l (J, Q)$$

where

$$J \triangleq \begin{bmatrix} K_{\bullet} & \tilde{V}_{\bullet}^{-1} \\ V_{\bullet}^{-1} & -V_{\bullet}^{-1} N \end{bmatrix}. \qquad \text{(2-23)}$$

Therefore, it is concluded that every stabilizing controller can be represented as a transfer function matrix $K = \mathbf{F}_l (J, Q)$, as in Figure 2-5, for some parameter matrix $Q \in RH_\infty$ (constrained only to be stable and proper and to make K proper). Using results of the above lemma, the affine parameterization of the closed-loop transfer matrix follows immediately.
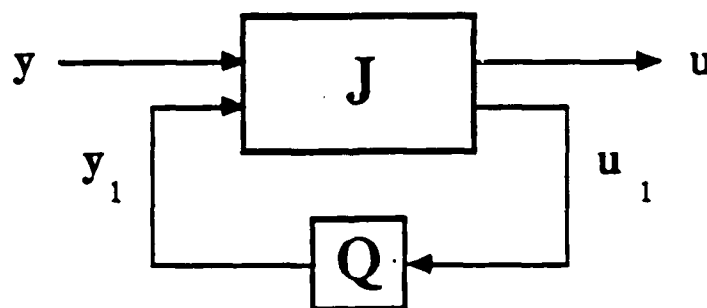
Figure 2-5.

Theorem 3

The set of all closed-loop transfer function matrices from v to e (Figure 2-2) achievable by an internally stabilizing proper controller as in Theorem 2 is

$$\left\{ T_{11} + T_{12}QT_{21} : Q \in RH_\infty, \quad I + D_{22}Q(\infty) \text{ invertible} \right\}.$$

where

$$T_{11} = P_{11} + P_{12}U_{\bullet}\tilde{M}P_{21} \tag{2-24a}$$

$$T_{12} = P_{12}M \tag{2-24b}$$

and

$$T_{21} = \tilde{M}P_{21} \tag{2-24c}$$

(P_{ij}, $\tilde{M}$, and $V_{\bullet}$ are defined as in Eqs. 2-2, 2-18, and 2-21 respectively).

Proof

Substituting Eq. 2-22a into $\mathbf{F}_l(P,K)$ in Eq. 2-1 yields

$$\mathbf{F}_l(P,K) = P_{11} + P_{12}(U_{\bullet} + MQ)(V_{\bullet} + NQ)^{-1}\left[I - \tilde{M}^{-1}\tilde{N}(U_{\bullet} + MQ)(V_{\bullet} + NQ)^{-1}\right]^{-1}P_{21}$$

$$= P_{11} + P_{12}(U_{\bullet} + MQ)\left[\tilde{M}(V_{\bullet} + NQ) - \tilde{N}(U_{\bullet} + MQ)\right]^{-1}\tilde{M}P_{21}$$

Eqs. 2-18 and 2-20 then yield $\tilde{M}N = \tilde{N}M$ and $\tilde{M}V_{\bullet} - \tilde{N}U_{\bullet} = I$. Therefore,

$$\mathbf{F}_l(P,K) = P_{11} + P_{12}(U_{\bullet} + MQ)\tilde{M}P_{21}$$

$$= (P_{11} + P_{12}U_{\bullet}\tilde{M}P_{21}) + (P_{12}M)Q(\tilde{M}P_{21})$$

$$= T_{11} + T_{12}QT_{21} \tag{2-25}$$

where T_{11}, T_{12}, and T_{21} are defined as in Eq. 2-24.

<div align="right">QED</div>

Remark

Eq. 2-25 is also a linear fractional transformation, $F_l(T,Q)$ with $T_{22} = 0$.

In the following, the coprime factorizations will be given in terms of state-space representations. Using these formula, an explicit realization of one choice of the interconnection matrix J (Eq. 2-23) is then derived.

For the next two lemmas, it is assumed that $G = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in R_p^{p \times m}$ where the pair (A,B) is stabilizable and the pair (C,A) is detectable. The following lemma shows that an rcf of G can be obtained using a stabilizing state feedback gain matrix F. The existence of F is guaranteed since (A,B) is stabilizable.

Lemma 4

A stabilizing state feedback F yields $rcf \quad G = NM^{-1}$ where

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array} \right] . \tag{2-26}$$

Note that for any nonsingular Z

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & BZ \\ \hline F & Z \\ C+DF & DZ \end{array} \right] . \tag{2-27}$$

is also a realization of an rcf of G.

By duality, to get an lcf of G, take H such that $A+HC$ is a stability matrix. Call such a stabilizing matrix H an output injection gain.

Lemma 4'

A stabilizing output injection gain H yields $lcf \quad G = \tilde{M}^{-1}\tilde{N}$ where

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} = \left[\begin{array}{c|cc} A+HC & H & B+HD \\ \hline C & I & D \end{array} \right] . \tag{2-28}$$

The next step is to specify U_o, V_o, $\tilde{U}_o$, $\tilde{V}_o$ to satisfy the doubly Bezout identity in Eq. 2-20. The idea behind the choice of these matrices is as follows. Using observer theory [Kai], find a controller K_o, and the associated matrices F, H achieving internal stability and then perform factorizations on P_{22} and K_o analogous to the ones just performed on G (Eqs. 2-26 and 2-28). Then the result summarized as in Lemma 1 implies that the left-hand side of Eq. 2-20 must be invertible in RH_∞. In addition, Eq. 2-20 is satisfied.

The transfer matrices N, M, $\tilde{N}$, and $\tilde{M}$ in Eq. 2-18 have the following realizations,

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline F & I \\ C_2+D_{22}F & D \end{array} \right] , \tag{2-29a}$$

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} = \left[\begin{array}{c|cc} A+HC_2 & H & B_2+HD_{22} \\ \hline C_2 & I & D_{22} \end{array} \right] . \tag{2-29b}$$

The realization equations for K_o are

$$\dot{z} = A\dot{z} + B_2u + H(C_2\dot{z} + D_{22}u - y)$$
$$u = F\dot{z} ,$$

that is,

$$K_o = \left[\begin{array}{c|c} A+B_2F+HC_2+HD_{22}F & -H \\ \hline F & 0 \end{array} \right] . \tag{2-30}$$

Define

$$\hat{A} = A+B_2F+HC_2+HD_{22}F , \qquad \hat{B} = -H ,$$

$$\hat{C} = F , \qquad\qquad\qquad\qquad \hat{D} = 0 ,$$

$$\hat{F} = C_2+D_{22}F , \qquad\qquad \hat{H} = -(B_2+HD_{22}) .$$

Following Eqs. 2-26 and 2-28, define

$$\begin{bmatrix} V_o \\ U_o \end{bmatrix} = \left[\begin{array}{c|c} \hat{A} + \hat{B}\hat{F} & \hat{B} \\ \hline \hat{F} & I \\ \hat{C} + \hat{D}\hat{F} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} A + BF & -H \\ C + DF & I \\ F & 0 \end{array}\right] \tag{2-31a}$$

and

$$\begin{bmatrix} \bar{V}_o & \bar{U}_o \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} + \hat{H}\hat{C} & \hat{H} & \hat{B} + \hat{H}\hat{D} \\ \hline \hat{C} & I & \hat{D} \end{array}\right] = \left[\begin{array}{c|cc} A + HC_2 & -(B_2 + HD_{22}) & -H \\ \hline F & I & 0 \end{array}\right]. \tag{2-31b}$$

Therefore,

$$\begin{bmatrix} M & U_o \\ N & V_o \end{bmatrix} = \left[\begin{array}{c|cc} A + B_2 F & B_2 & -H \\ \hline F & I & 0 \\ C_2 + D_{22}F & D_{22} & I \end{array}\right]. \tag{2-32a}$$

$$\begin{bmatrix} \bar{V}_o & -\bar{U}_o \\ -\bar{N} & \bar{M} \end{bmatrix} = \left[\begin{array}{c|cc} A + HC_2 & -(B_2 + HD_{22}) & H \\ \hline F & I & 0 \\ C_2 & -D_{22} & I \end{array}\right]. \tag{2-32b}$$

and Eq. 2-20 is satisfied.

A realization of J is now immediate. Substitution of Eqs. 2-29a, 2-30, and 2-31 into Eq. 2-23 leads to, after simplification,

$$J = \left[\begin{array}{c|cc} A + B_2 F + HC_2 + HD_{22}F & -H & B_2 + HD_{22} \\ \hline F & 0 & I \\ -(C_2 + D_{22}F) & I & -D_{22} \end{array}\right]. \tag{2-33}$$

Theorem 2 provides a parametrization, in terms of Q, of all proper Ks which achieve internal stability in Figure 2-1. Substitution of the transfer function relationship of the block diagram in Figure 2-5 into that in Figure 2-1 leads to the one in Figure 2-6. Elimination of the signals u and y leads to the transfer function relationship diagrammed as Figure 2-7 for a suitable transfer matrix T. Thus all closed-loop transfer matrices are representable as in Figure 2-7.
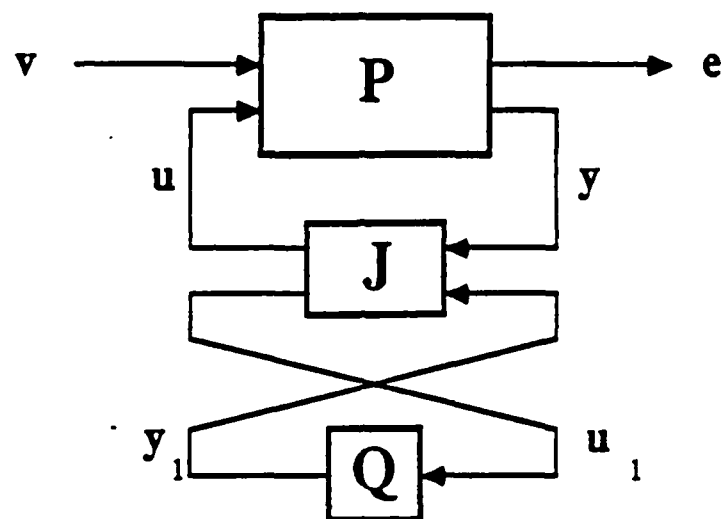
Figure 2-6.
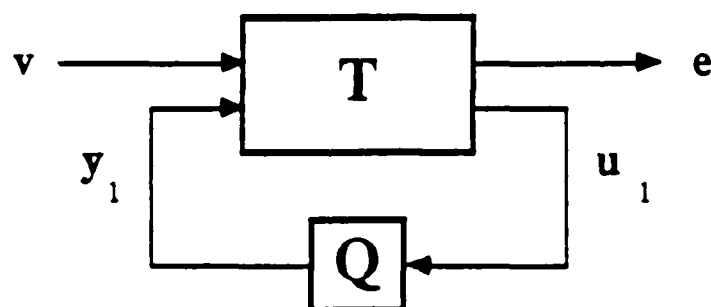
Figure 2-7.

The following theorem gives one particular realization of T.

Theorem 5

Consider T as in Figure 2-7, then T has a representation in the form

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & 0 \end{bmatrix} = \left[\begin{array}{cc|cc} A+B_2F & -HC_2 & -HD_{21} & B_2 \\ 0 & A+HC_2 & B_1+HD_{21} & 0 \\ \hline C_1+D_{12}F & C_1 & D_{11} & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array} \right] \tag{2-34}$$

where the matrices F and H are defined as in Eqs. 2-29a and 2-29b.

Proof

(i) T_{12}:

Substituting Eqs. 2-8 and 2-29a into Eq. 2-24b,

$$T_{12} = P_{12}M$$

$$= \left[\begin{array}{c|c} A & B_2 \\ \hline C_1 & D_{12} \end{array} \right] * \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline F & I \end{array} \right]$$

$$= \left[\begin{array}{cc|c} A & B_2F & B_2 \\ 0 & A+B_2F & B_2 \\ \hline C_1 & D_{12}F & D_{12} \end{array} \right]$$

Applying the similarity transformation $\begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}$ yields,

$$T_{12} = \left[\begin{array}{cc|c} A & 0 & 0 \\ 0 & A+B_2F & B_2 \\ \hline C_1 & C_1+D_{12}F & D_{12} \end{array} \right] = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array} \right] \tag{2-35a}$$

(ii) T_{21}:

Substituting Eqs. 2-8 and 2-29b into Eq. 2-24c,

$$T_{21} = \bar{M}P_{21}$$

$$= \left[\begin{array}{c|c} A+HC_2 & H \\ \hline C_2 & I \end{array} \right] * \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & D_{21} \end{array} \right]$$

$$= \left[\begin{array}{cc|c} A + HC_2 & HC_2 & HD_{21} \\ 0 & A & B_1 \\ \hline C_2 & C_2 & D_{21} \end{array}\right]$$

Applying the similarity transformation $\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$ yields,

$$T_{21} = \left[\begin{array}{cc|c} A + HC_2 & 0 & B_1 + HD_{21} \\ 0 & A & B_1 \\ \hline C_2 & 0 & D_{12} \end{array}\right] = \left[\begin{array}{c|c} A + HC_2 & B_1 + HD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \qquad (2\text{-}35b)$$

(iii) T_{11}:

From Eqs. 2-8, 2-31a, and 2-35b,

$$P_{12} U_\bullet \bar{M} P_{21} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_1 & D_{12} \end{array}\right] * \left[\begin{array}{c|c} A + B_2 F & -H \\ \hline F & 0 \end{array}\right] * \left[\begin{array}{c|c} A + HC_2 & B_1 + HD_{21} \\ \hline C_2 & D_{21} \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} A & B_2 F & 0 & 0 \\ 0 & A + B_2 F & -HC_2 & -HD_{21} \\ 0 & 0 & A + HC_2 & B_1 + HD_{21} \\ \hline C_1 & D_{12} F & 0 & 0 \end{array}\right]$$

Applying the similarity transformation $\begin{bmatrix} I & -I & -I \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$ yields,

$$P_{12} U_\bullet \bar{M} P_{21} = \left[\begin{array}{ccc|c} A & 0 & 0 & -B_1 \\ 0 & A + B_2 F & -HC_2 & -HD_{21} \\ 0 & 0 & A + HC_2 & B_1 + HD_{21} \\ \hline C_1 & C_1 + D_{12} F & C_1 & 0 \end{array}\right]$$

Therefore, from Eq. 2-24a,

$$T_{11} = P_{11} + P_{12} U_\bullet \bar{M} P_{21}$$

$$= \left[\begin{array}{c|c} A & B_1 \\ \hline C_1 & D_{11} \end{array}\right] + \left[\begin{array}{ccc|c} A & 0 & 0 & -B_1 \\ 0 & A + B_2 F & -HC_2 & -HD_{21} \\ 0 & 0 & A + HC_2 & B_1 + HD_{21} \\ \hline C_1 & C_1 + D_{12} F & C_1 & 0 \end{array}\right]$$

$$
= \left[\begin{array}{cc|c} A + B_2 F & -HC_2 & -HD_{21} \\ 0 & A + HC_2 & B_1 + HD_{21} \\ \hline C_1 + D_{12}F & C_1 & D_{11} \end{array} \right] \qquad \text{(2-35c)}
$$

Combining Eqs. 2-35a through 2-35c and the fact that $T_{22} = 0$, Eq. 2-34 follows immediately.

<div align="right">QED</div>

Remark

The following expression is an alternative realization for T which was given by Doyle [D16].

$$
T = \left[\begin{array}{cc|cc} A + B_2 F & -B_2 F & B_1 & B_2 \\ 0 & A + HC_2 & B_1 + HD_{21} & 0 \\ \hline C_1 + D_{12}F & -D_{12}F & D_{11} & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array} \right] \qquad \text{(2-36)}
$$

It is important to note that the closed-loop transfer matrix is simply an "affine" function of the controller parameter matrix Q and that the coefficient matrices T_{ij} have very simple realizations; namely as described by Eq. 2-34 (or Eq. 2-36).

Note that the Youla parametrization and associated observer-based controller described above allow choice of the matrices F and H. In the next section, specific choices of F and H will be presented such that the affine parametrization in Theorem 3 has additional properties, namely that, T_{12} and T_{21} are inner and co-inner respectively. The algebraic Riccati equation will play an essential role in obtaining such a parametrization.

2.3 Coprime Factorisation with Inner Numerator

Section 2.2 contains a summary of methods for finding the interconnection matrix J so that the substitution $K = \mathbf{F}_l(J, Q)$ yields

$$
\begin{aligned}
\mathbf{F}_l(P, K) &= \mathbf{F}_l(P, \mathbf{F}_l(J, Q)) \\
&= \mathbf{F}_l(T, Q) \\
&= T_{11} + T_{12} Q T_{21}
\end{aligned}
$$

with the additional requirement that $T \in H_\infty$ and

$$\mathbf{F}_l(P, K) \quad \text{internally stable}$$

$$\text{if and only if} \quad Q \in H_\infty .$$

This parametrizes all stabilizing controllers K in terms of a stability matrix $Q \in H_\infty$ in addition to providing an affine parametrization of all stable $\mathbf{F}_l(P, K)$. The actual structure of J was derived in terms of an observer-based compensator. The stabilizing state feedback and output injection of the observer-based compensator were shown to provide coprime factorizations of P_{22} and solve the Bezout identities necessary to provide the parametrization of all stabilizing controllers.

In this section, the requirement is added that T_{12} and T_{21} be inner and co-inner respectively; that is, $T_{12}^* T_{12} = I$ and $T_{21} T_{21}^* = I$. In addition, $T_\perp$ and $\tilde{T}_\perp$ will be found so that $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}$ are square and inner.

The key idea behind the factorization in this section is the connection between inner functions, algebraic Riccati equations (AREs), and spectral factorizations. The stabilizing solution of the ARE's will be needed in order to construct the desired factorizations.

Coprime factorizations with inner numerator and complementary inner factors (CIF) will be obtained which involve using a state feedback or output injection gain based on the stabilizing solution of a particular ARE. This provides a reliable computational method based on standard approaches to finding solutions of AREs [C5,K6,L1,M1,M2,M3,P3,V1]. A complete treatment on this subject can be found elsewhere [A3,A4,C2,D16,W1,Y1].

Consider the Algebraic Riccati Equation,

$$E^T X + XE - XWX + Q = 0 \qquad \text{(ARE)}$$

with the associated Hamiltonian matrix

$$A_H = \begin{bmatrix} E & -W \\ -Q & -E^T \end{bmatrix}, \qquad \text{(Hamiltonian)}$$

where

$$E, W, Q \in \mathbf{R}^{n \times n}, \quad W = W^T \geq 0, \quad \text{and} \quad Q = Q^T.$$

The following theorem and corollary characterize the relationship between spectral factorization, AREs, and decomposition of Hamiltonians. The proofs can be found in [D16].

Theorem 1

Let $A, B, P, S,$ and R be matrices of compatible dimensions for the ARE such that $P = P^T$, $R = R^T > 0$, with the pair (A, B) stabilizable and the pair (P, A) detectable. Then the following statements are equivalent:

(i) The para-Hermitian rational matrix

$$\Gamma(s) = \begin{bmatrix} B^T(-sI - A^T)^{-1} & I \end{bmatrix} \begin{bmatrix} P & S \\ S^T & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}$$

satisfies

$$\Gamma(j\omega) > 0 \quad \text{for all} \quad 0 \leq \omega \leq \infty.$$

(ii) For $E = A - BR^{-1}S^T$, $W = BR^{-1}B^T$, and $Q = P - SR^{-1}S^T$, there exists a unique real matrix $X = X^T$ such that

$$E^T X + XE - XWX + Q = 0$$

and $(E - BR^{-1}B^T X)$ is a stability matrix.

(iii)   The Hamiltonian matrix

$$A_H = \begin{bmatrix} A - BR^{-1}S^T & -BR^{-1}B^T \\ -P + SR^{-1}S^T & -(A - BR^{-1}S^T)^T \end{bmatrix}$$

has no eigenvalues on the $j\omega$-axis.

Remark

The unique stabilizing solution of the ARE (X as in (ii) of Theorem 1) will be denoted by $\mathbf{Ric}\,(A_H)$ throughout where A_H is the associated Hamiltonian matrix.

Corollary 2

If the conditions in Theorem 1 are satisfied then $\exists\ M \in R_p$ such that $M^{-1} \in RH_\infty$ and

$$\Gamma = M^*RM .$$

A particular realization of one such M is

$$M = \begin{bmatrix} A & B \\ \hline -F & I \end{bmatrix} \tag{2-37}$$

where $F = -R^{-1}(S^T + B^TX)$.

In the following, the special form of coprime factorizations required to simplify the general H_α ($\alpha = 2, \infty$) optimal control to a distance problem will be developed. In particular, explicit realizations are given for coprime factorizations $G = NM^{-1}$ with inner numerator N (Theorem 4) and for the complementary inner factor $N_\perp$ which completes the inner numerator to make $\begin{bmatrix} N & N_\perp \end{bmatrix}$ square and inner (Theorem 5). The results will be stated as theorems for rcf s. The duals for lcf s following just as for the general case of coprime factorization developed Section 2.2.

The following lemma summarizes the necessary and sufficient conditions of an inner transfer matrix with state-space representation. The proof can be found in [D16].

Lemma 3

The transfer matrix $N = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \in RH_\infty$ is inner if and only if

(i) $\hat{B}^T X + \hat{D}^T \hat{C} = 0$ (2-38)

and

(ii) $\hat{D}^T \hat{D} = I$ (2-39)

where the observability gramian X solves

$$\hat{A}^T X + X\hat{A} + \hat{C}^T \hat{C} = 0 \qquad (2\text{-}40)$$

For the next two results, it is assumed that $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in R_p^{\prime \times m}$ and the realization for G is minimal. The notation $R_D^{1/2}$ $(R_D > 0)$ will be used to denote the square root matrix such that $(R_D^{1/2})^T R_D^{1/2} = R_D$ (or $R_D^{1/2}(R_D^{1/2})^T = R_D$) and use "$D_\perp$" for any orthogonal complement of D so that $\left[DR_D^{-1/2} \ D_\perp\right]$ (with $R_D = D^T D$) is square and orthogonal.

From Lemma 2.2.4, a stabilizing state feedback F yields rcf $G = NM^{-1}$ where

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & BZ \\ \hline F & Z \\ C+DF & DZ \end{array}\right]. \qquad (2\text{-}41)$$

and Z can be any nonsingular matrix. To obtain an rcf with N inner, we simply need to use Eqs. 2-38 through 2-41 to solve for F and Z. This yields the following result.

Theorem 4

Assume $p \geq m$. Then, there exists an rcf $G = NM^{-1}$ with N inner if and only if $G^*G > 0$ on the $j\omega$-axis, including at ∞. This factorization is unique up to a constant orthogonal multiple. A particular realization for the factorization is

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A + BF & BR_D^{-1/2} \\ \hline F & R_D^{-1/2} \\ C + DF & DR_D^{-1/2} \end{array} \right] \quad \in \quad RH_\infty^{(m+p) \times m} \qquad (2\text{-}42)$$

where

$$R_D = D^T D > 0$$

$$F = -R_D^{-1}(B^T X + D^T C) \qquad (2\text{-}43)$$

and

$$X = \mathbf{Ric} \left[\begin{array}{cc} A - BR_D^{-1}D^T C & -BR_D^{-1}B^T \\ -C^T D_\perp D_\perp^{\ T} C & -(A - BR_D^{-1}D^T C)^T \end{array} \right] \geq 0 \qquad (2\text{-}44)$$

In a similar manner Eqs. 2-38 through 2-40 can be used to obtain the complementary inner factor (CIF) in the following theorem.

Theorem 5

If $p > m$ in Theorem 4, then there exists a complementary inner factor (CIF) $N_\perp \in RH_\infty^{p \times (p-m)}$ such that the matrix $\begin{bmatrix} N & N_\perp \end{bmatrix}$ is square and inner. A particular realization is

$$N_\perp = \left[\begin{array}{c|c} A + BF & -X^\dagger C^T D_\perp \\ \hline C + DF & D_\perp \end{array} \right] \qquad (2\text{-}45)$$

where X and F are from Theorem 4 and $X^\dagger$ is the pseudo-inverse of X .

Remarks

(1)   If $G \in RH_\infty^{p \times m}$ in Theorem 4, then M is a "unit" in RH_∞ and M^{-1} is "outer". In this case, $G = N(M^{-1})$ is called an "inner-outer factorization" (IOF).

(2)   Dual results for all factorizations can be obtained when $p \leq m$. In these factorizations, output injection using the dual Riccati solution replaces state feedback to obtain corresponding left factorizations.

In the rest of this section, the results of the Youla parametrization which parameterizes all stabilizing controllers and coprime factorization with inner numerator will be combined in a way that is convenient for solving H_2-optimal and H_∞-optimal control problems.

Assume P as the transfer function matrix associated with Eqs. 2-2 and 2-33 and neither P_{12} nor P_{21} has transmission zeros (see reference [K1] for definition of transmission zeros) on the $j\omega$-axis (including ∞). This implies that

$$R_D = D_{12}^T D_{12} > 0 \quad \text{and} \quad \tilde{R}_D = D_{21} D_{21}^T > 0 .$$

Under these assumptions, let $D_\perp = (D_{12})_\perp$ and $\tilde{D}_\perp = (D_{21})_\perp$, that is, $\left[D_{12} R_D^{-1/2} \ D_\perp \right]$ and $\begin{bmatrix} \tilde{R}_D^{-1/2} D_{21} \\ \tilde{D}_\perp \end{bmatrix}$ are orthogonal matrices where

$$(R_D^{1/2})^T R_D^{1/2} = R_D \quad \text{and} \quad \tilde{R}_D^{1/2} (\tilde{R}_D^{1/2})^T = \tilde{R}_D .$$

By Theorem 4, $N_{12} \triangleq T_{12} R_D^{-1/2}$ is inner where the state feedback gain matrix F is

$$F = -R_D^{-1}(D_{12}^T C_1 + B_2^T X) \tag{2-46a}$$

and

$$X = \text{Ric} \begin{bmatrix} A - B_2 R_D^{-1} D_{12}^T C_1 & -B_2 R_D^{-1} B_2^T \\ -C_1^T D_\perp D_\perp^T C_1 & -(A - B_2 R_D^{-1} D_{12}^T C_1)^T \end{bmatrix} . \tag{2-46b}$$

Similarly, using the dual form of Theorem 4, $N_{21} \triangleq \tilde{R}_D^{-1/2} T_{21}$ is co-inner where the output-injection gain H is

$$H = -(B_1 D_{21}^T + Y C_2^T) \tilde{R}_D^{-1} \tag{2-47a}$$

and

$$Y = \text{Ric} \begin{bmatrix} (A - B_1 D_{21}^T \tilde{R}_D^{-1} C_2)^T & -C_2^T \tilde{R}_D^{-1} C_2 \\ -B_1 \tilde{D}_\perp^T \tilde{D}_\perp B_1^T & -(A - B_1 D_{21}^T \tilde{R}_D^{-1} C_2) \end{bmatrix} . \tag{2-47b}$$

Then, $N_{12}^* N_{12} = I$ and $N_{21} N_{21}^* = I$. Also, let $N_\perp$ and $\tilde{N}_\perp$ be the corresponding CIFs so that

$$\begin{bmatrix} N_{12} & N_\perp \end{bmatrix} = \left[\begin{array}{c|cc} A + B_2 F & B_2 R_D^{-1/2} & -X^\dagger C_1^\dagger D_\perp \\ \hline C_1 + D_{12} F & D_{12} R_D^{-1/2} & D_\perp \end{array} \right] \tag{2-48}$$

$$\begin{bmatrix} N_{21} \\ \tilde{N}_\perp \end{bmatrix} = \left[\begin{array}{c|c} A + H C_2 & B_1 + H D_{21} \\ \hline \tilde{R}_D^{-1/2} C_2 & \tilde{R}_D^{-1/2} D_{21} \\ -\tilde{D}_\perp B_1^\dagger Y^\dagger & \tilde{D}_\perp \end{array} \right] \tag{2-49}$$

Therefore,

$$\begin{aligned} \mathbf{F}_l(P,K) &= \mathbf{F}_l(T,Q) \\ &= T_{11} + T_{12} Q T_{21} \\ &= T_{11} - N_{12}(-R_D^{1/2} Q \tilde{R}_D^{1/2}) N_{21} \\ &= T_{11} - N_{12} \hat{Q} N_{21} \end{aligned} \tag{2-50}$$

where

$$\hat{Q} = -R_D^{1/2} Q \tilde{R}_D^{1/2} \in RH_\infty \tag{2-51}$$

is the new stable parameter matrix.

Because both the $|| \cdot ||_2$ and $|| \cdot ||_\infty$ norms are unitarily invariant, an alternative expression is possible. For any $\hat{Q} \in RH_\infty$,

$$\begin{aligned} &\left|\left| T_{11} - N_{12} \hat{Q} N_{21} \right|\right|_\sigma \\ &= \left|\left| \begin{bmatrix} N_{12} & N_\perp \end{bmatrix}^* (T_{11} - N_{12} \hat{Q} N_{21}) \begin{bmatrix} N_{21} \\ \tilde{N}_\perp \end{bmatrix}^* \right|\right|_\sigma \\ &= \left|\left| \begin{bmatrix} N_{12}^* T_{11} N_{21}^* - \hat{Q} & N_{12}^* T_{11} \tilde{N}_\perp^* \\ N_\perp^* T_{11} N_{21}^* & N_\perp^* T_{11} \tilde{N}_\perp^* \end{bmatrix} \right|\right|_\sigma \\ &= \left|\left| \begin{bmatrix} R_{11} - \hat{Q} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right|\right|_\sigma \end{aligned} \tag{2-52}$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} T_{11} \begin{bmatrix} N_{21}^* & \tilde{N}_\perp^* \end{bmatrix}. \tag{2-53}$$

Now, the $\alpha = 2$ case is particularly simple. Since

$$\left\| \begin{bmatrix} R_{11} - \hat{Q} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2 = \left(\left\| R_{11} - \hat{Q} \right\|_2^2 + \left\| \begin{bmatrix} 0 & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2^2 \right)^{1/2}$$

the optimal $\hat{Q}$ is found immediately to be

$$\hat{Q}_{opt} = \mathbf{P}_{\mathbf{H}_2}(R_{11}) \tag{2-54}$$

and the the corresponding optimal Q is

$$Q_{opt} = -R_D^{-1/2}\mathbf{P}_{\mathbf{H}_2}(R_{11})\bar{R}_D^{-1/2} \tag{2-55}$$

The case of $\alpha = \infty$ is more complicated and the formulation in Eq. 2-52 is called the "general distance problem" (in short, **GDP**) which will be investigated in great detail in the next three chapters.

A state-space realization for R is given in the following theorem. It is suggested that all the poles of R are in the open rhp, i.e., R is completely unstable. For convenience, R is represented as the cascade operation from two system matrices.

Theorem 6

Consider R from the formulation in Eq. 2-53, then R has the realization

$$R = \begin{bmatrix} -(A + B_2 F)^T & (C_1 + D_{12}F)^T & -XH \\ -(B_2 R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T & 0 \\ -D_\perp^T C_1 X^\dagger & D_\perp^T & 0 \end{bmatrix} * \begin{bmatrix} -(A + HC_2)^T & -(\bar{R}_D^{-1/2}C_2)^T & Y^\dagger B_1 \bar{D}_\perp^T \\ C_1 Y + D_{11}(B_1 + HD_{21})^T & D_{11}(\bar{R}_D^{-1/2}D_{21})^T & D_{11}\bar{D}_\perp^T \\ 0 & \bar{R}_D^{1/2} & 0 \end{bmatrix}.$$

Proof

Note that the realization of T_{11}, Eq. 2-35c, can be expressed as the cascade of two systems, i.e.,

$$T_{11} = \begin{bmatrix} A + B_2 F & -HC_2 & -HD_{21} \\ 0 & A + HC_2 & B_1 + HD_{21} \\ \hline C_1 + D_{12}F & C_1 & D_{11} \end{bmatrix}$$

$$= \begin{bmatrix} A + B_2 F & 0 & -H \\ \hline C_1 + D_{12}F & I & 0 \end{bmatrix} * \begin{bmatrix} A + HC_2 & B_1 + HD_{21} \\ \hline C_1 & D_{11} \\ C_2 & D_{21} \end{bmatrix}$$

62

$$= (T_{11})_A (T_{11})_B$$

It is convenient to compute $\begin{bmatrix} N_{12}^* \\ N_{\perp}^* \end{bmatrix} (T_{11})_A$ and $(T_{11})_B \begin{bmatrix} N_{21}^* & \tilde{N}_{\perp}^* \end{bmatrix}$ separately.

(a)

$$\begin{bmatrix} N_{12}^* \\ N_{\perp}^* \end{bmatrix} (T_{11})_A = \left[\begin{array}{c|c} -(A+B_2F)^T & (C_1+D_{12}F)^T \\ \hline -(B_2R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T \\ D_{\perp}^T C_1 X^{\dagger} & D_{\perp}^T \end{array} \right] * \left[\begin{array}{c|cc} A+B_2F & 0 & -H \\ \hline C_1+D_{12}F & I & 0 \end{array} \right]$$

$$= \left[\begin{array}{cc|cc} -(A+B_2F)^T & (C_1+D_{12}F)^T(C_1+D_{12}F) & (C_1+D_{12}F)^T & 0 \\ 0 & A+B_2F & 0 & -H \\ \hline -(B_2R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T(C_1+D_{12}F) & (D_{12}R_D^{-1/2})^T & 0 \\ D_{\perp}^T C_1 X^{\dagger} & D_{\perp}^T(C_1+D_{12}F) & D_{\perp}^T & 0 \end{array} \right]$$

$$= \left[\begin{array}{cc|cc} -(A+B_2F)^T & C_1D_{\perp}D_{\perp}^TC_1 + XB_2R_D^{-1}B_2X & (C_1+D_{12}F)^T & 0 \\ 0 & A+B_2F & 0 & -H \\ \hline -(B_2R_D^{-1/2})^T & -(B_2R_D^{-1/2})^T X & (D_{12}R_D^{-1/2})^T & 0 \\ D_{\perp}^T C_1 X^{\dagger} & D_{\perp}^T C_1 & D_{\perp}^T & 0 \end{array} \right]$$

Applying the similarity transformation $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$ (X is as in Eq. 2-46a) yields,

$$\begin{bmatrix} N_{12}^* \\ N_{\perp}^* \end{bmatrix} (T_{11})_A = \left[\begin{array}{cc|cc} -(A+B_2F)^T & (*) & (C_1+D_{12}F)^T & -XH \\ 0 & A+B_2F & -X(C_1+D_{12}F)^T & -H \\ \hline -(B_2R_D^{-1/2})^T & 0 & (D_{12}R_D^{-1/2})^T & 0 \\ D_{\perp}^T C_1 X^{\dagger} & D_{\perp}^T C_1(I-XX^{\dagger}) & D_{\perp}^T & 0 \end{array} \right]$$

where

$$(*) = (A+B_2F)^T X + X(A+B_2F) + XB_2R_D^{-1}B_2^TX + C_1D_{\perp}D_{\perp}^TC_1$$

$$= (A-B_2D_{12}^TC_1)^T X + X(A-B_2D_{12}^TC_1) - XB_2R_D^{-1}B_2^TX + C_1D_{\perp}D_{\perp}^TC_1$$

$$= 0 .$$

Also,

$$\mathbf{Ker}(X) \subset \mathbf{Ker}(D_\perp^T C_1) \implies D_\perp^T C_1(I - XX^\dagger) = 0 \ .$$

$$\therefore \quad \begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} (T_{11})_A = \begin{bmatrix} -(A + B_2 F)^T & (C_1 + D_{12}F)^T & -XH \\ \hline -(B_2 R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T & 0 \\ D_\perp^T C_1 X^\dagger & D_\perp^T & 0 \end{bmatrix} \ . \tag{2-56}$$

(b)

$$(T_{11})_B \begin{bmatrix} N_{21}^* & \tilde{N}_\perp^* \end{bmatrix} = \begin{bmatrix} A + HC_2 & B_1 + HD_{21} \\ \hline C_1 & D_{11} \\ C_2 & D_{21} \end{bmatrix} * \begin{bmatrix} -(A + HC_2)^T & (\tilde{R}_D^{-1/2}C_2)^T & -Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline -(B_1 + HD_{21})^T & (\tilde{R}_D^{-1/2}D_{21})^T & \tilde{D}_\perp^T \end{bmatrix}$$

$$= \begin{bmatrix} A + HC_2 & -(B_1 + HD_{21})(B_1 + HD_{21})^T & (B_1 + HD_{21})D_{21}^T & (B_1 + HD_{21})\tilde{D}_\perp^T \\ 0 & -(A + HD_{21})^T & (\tilde{R}_D^{-1/2}C_2)^T & -Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline C_1 & -D_{11}(B_1 + HD_{21})^T & D_{11}(\tilde{R}_D^{-1/2}D_{21})^T & D_{11}\tilde{D}_\perp^T \\ C_2 & D_{21}(B_1 + HD_{21})^T & \tilde{R}_D^{1/2} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} A + HC_2 & -(B_1 \tilde{D}_\perp^T \tilde{D}_\perp B_1^T) & -YC_2^T & B_1 \tilde{D}_\perp^T \\ 0 & -(A + HC_2)^T & (\tilde{R}_D^{-1/2}C_2)^T & -Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline C_1 & -D_{11}(B_1 + HD_{21})^T & D_{11}(\tilde{R}_D^{-1/2}D_{21})^T & D_{11}\tilde{D}_\perp^T \\ C_2 & -C_2 Y & \tilde{R}_D^{1/2} & 0 \end{bmatrix} \ .$$

Applying the similarity transformation $\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$ (Y is as in Eq. 2-47a) yields,

$$(T_{11})_B \begin{bmatrix} N_{21}^* & \tilde{N}_\perp^* \end{bmatrix} = \begin{bmatrix} A + HC_2 & (\Theta) & 0 & (I - YY^\dagger)B_1 \tilde{D}_\perp^T \\ 0 & -(A + HC_2)^T & (\tilde{R}_D^{-1/2}C_2)^T & -Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline C_1 & -C_1 Y - D_{11}(B_1 + HD_{21})^T & D_{11}(\tilde{R}_D^{-1/2}D_{21})^T & D_{11}\tilde{D}_\perp^T \\ C_2 & 0 & \tilde{R}_D^{1/2} & 0 \end{bmatrix}$$

where

$$(\Theta) = -\left\{ (A + HC_2)Y + Y(A + HC_2)^T + B_1 \tilde{D}_\perp^T \tilde{D}_\perp B_1^T + YC_2^T \tilde{R}_D^{-1} C_2 Y \right\}$$

$$= -\left\{ (A - B_1 D_{21}^T \tilde{R}_D^{-1} C_2)Y + Y(A - B_1 D_{21}^T \tilde{R}_D^{-1} C_2)^T - YC_2^T \tilde{R}_D^{-1} C_2 Y + B_1 \tilde{D}_\perp^T \tilde{D}_\perp B_1^T \right\}$$

$$= 0 \; .$$

Also,

$$\mathbf{Ker}(Y) \subset \mathbf{Ker}(\tilde{D}_\perp B_1^T) \quad \Longrightarrow \quad (I - YY^\dagger)B_1\tilde{D}_\perp^T = 0 \; .$$

$$\therefore \quad (T_{11})_B \begin{bmatrix} N_{21} & \tilde{N}_\perp^* \end{bmatrix} = \left[\begin{array}{c|cc} -(A + HC_2)^T & -(\tilde{R}_D^{-1/2}C_2)^T & Y^\dagger B_1\tilde{D}_\perp^T \\ \hline C_1Y + D_{11}(B_1 + HD_{21})^T & D_{11}(\tilde{R}_D^{-1/2}D_{21})^T & D_{11}\tilde{D}_\perp^T \\ 0 & \tilde{R}_D^{1/2} & 0 \end{array} \right] \; . \quad (2\text{-}57)$$

Combining Eqs. 2-56 and 2-57, the theorem follows immediately.

<div align="right">QED</div>

Remark

This realization is not unique, An equivalent expression can be found in [D16].

It is clear that R has all of its poles in the open rhp. Thus, the projection onto $H_2 + \mathbb{C}$ leaves only the constant term. Therefore, in the H_2 case one has the following result.

Corollary 7

The H_2-optimal solution is

$$Q_{opt} = -R_D^{-1}D_{12}^T D_{11}D_{21}^T\tilde{R}_D^{-1} \quad (2\text{-}58)$$

In the H_∞ case, the solution of the general distance problem is not trivial which will be discussed extensively in the next three chapters. To illustrate the approach described in this chapter, two examples, corresponding to H_2 and H_∞ optimization respectively, will be presented in the next section.

2.4 Examples

In this section, two examples will be presented. In the first example, the familiar linear quadratic Gaussian (LQG) formulation is treated using the general H_2 optimal control theory described in previous sections.

Example 1: (LQG)

Consider a linear time-invariant system with the following stochastic description :

$$\dot{z} \;=\; Az \;+\; Bu \;+\; Gd \tag{2-59a}$$

$$y \;=\; Cz \;+\; Nn \tag{2-59b}$$

where

$A,G \in \mathbf{R}^{n \times n}, \quad B \in \mathbf{R}^{n \times m}, \quad C \in \mathbf{R}^{p \times n} \text{ and } N \in \mathbf{R}^{p \times p},$

(A,B) is controllable and (C,A) is observable,

$\mathbf{E}[d(t)] = 0, \quad \mathbf{E}[d(t)d^T(\tau)] = I\delta(t-\tau),$

$\mathbf{E}[n(t)] = 0, \quad \mathbf{E}[n(t)n^T(\tau)] = I\delta(t-\tau),$

and

N is nonsingular.

The standard LQG formulation involves minimization of the (expectation of) quadratic cost function, that is, minimization of

$$\mathbf{E}\left[\int_0^\infty \{ \parallel z \parallel_Q^2 \;+\; \parallel u \parallel_W^2 \} \, dt\right] \tag{2-60}$$

with respect to z and u. In addition, the factorizations $Q = L^T L \geq 0$ and $W = M^T M > 0$ are assumed. A block diagram representation is shown in Figure 2-8 which can be further rearranged into the general synthesis framework suggested in Figure 2-9.
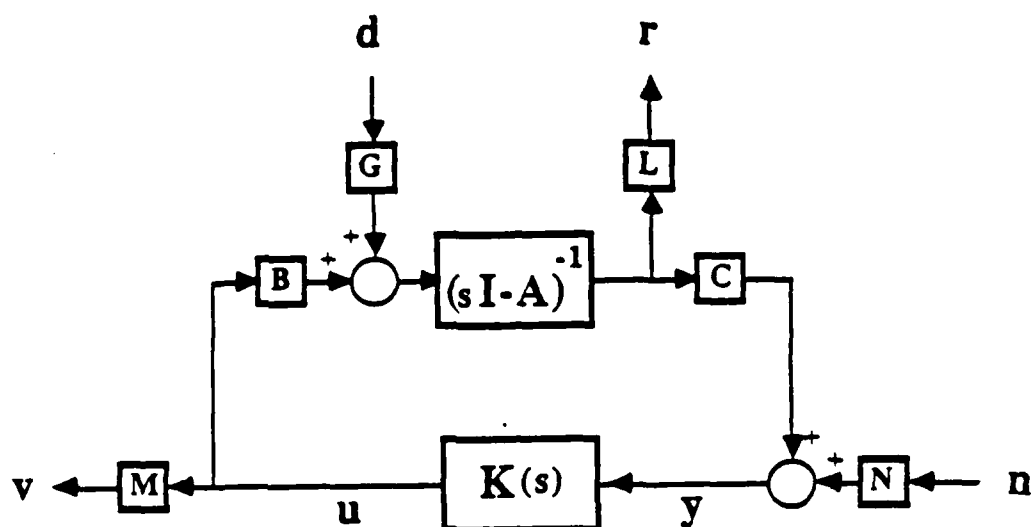
Figure 2-8. Block diagram of representation LQG
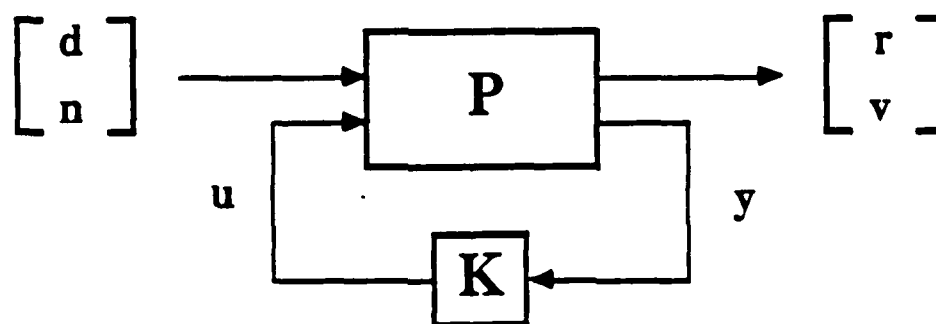
Figure 2-9. General synthesis framework for LQG

Let $\Phi(s) = (sI - A)^{-1}$

then

$$\begin{bmatrix} r \\ v \\ y \end{bmatrix} = \begin{bmatrix} L\Phi G & 0 & L\Phi B \\ 0 & 0 & M \\ C\Phi G & N & C\Phi B \end{bmatrix} \begin{bmatrix} d \\ n \\ u \end{bmatrix}.$$

The interconnection matrix P of Figure 2-9 in partitioned form is simply

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} L\Phi G & 0 & L\Phi B \\ 0 & 0 & M \\ C\Phi G & N & C\Phi B \end{bmatrix},$$

and has the following realization

$$P = \left[\begin{array}{c|ccc} A & G & 0 & B \\ \hline L & 0 & 0 & 0 \\ 0 & 0 & 0 & M \\ C & 0 & N & 0 \end{array} \right].$$

By assumptions, (A,B) is a controllable pair and (C,A) is an observable pair. Therefore, according to Section 2.3, the state feedback gain matrix F and observer gain matrix H can then be found by solving the appropriate AREs. Explicit solutions for F and H are:

(i) $\quad F = -(M^TM)^{-1}\left\{ \begin{bmatrix} 0 & M \end{bmatrix} \begin{bmatrix} L \\ 0 \end{bmatrix} + B^TX \right\} = -(M^TM)^{-1}B^TX$

$\qquad$ where $\quad X = \text{Ric} \begin{bmatrix} A & -B(M^TM)^{-1}B^T \\ -L^TL & -A^T \end{bmatrix}$

(ii) $\quad H = -\left\{ \begin{bmatrix} G & 0 \end{bmatrix} \begin{bmatrix} 0 \\ N^T \end{bmatrix} + YC^T \right\}(NN^T)^{-1} = -YC^T(NN^T)^{-1}$

$\qquad$ where $\quad Y = \text{Ric} \begin{bmatrix} A^T & -C^T(NN^T)^{-1}C \\ -GG^T & -A \end{bmatrix}$

Hence, from Eq. 2-32a, the matrix J has the following realization:

$$J = \left[\begin{array}{c|cc} A+BF+HC & -H & B \\ \hline F & 0 & I \\ -C & I & 0 \end{array}\right].$$

Since the external inputs are assumed to be white Gaussian, and the performance is measured in terms of the minimum-mean-square-error (MMSE) criterion, it can be shown that this is an H_2 optimal control problem. Therefore, from Corollary 2.3.7, the H_2-optimal solution is

$$Q_{opt} = -(M^TM)^{-1}\begin{bmatrix}0 & M^T\end{bmatrix}\begin{bmatrix}0 & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix}0 \\ N^T\end{bmatrix}(NN^T)^{-1} = 0.$$

The optimal stabilizing controller K_{opt} can be computed from the LFT $\mathbf{F}_l(J,Q_{opt})$ and has the following realization:

$$K_{opt} = \mathbf{F}_l(J,0) = J_{11} = \left[\begin{array}{c|c} A+BF+HC & -H \\ \hline F & 0 \end{array}\right]$$

It is not difficult to identify that the transfer function of optimal controller K_{opt} is exactly the same as that derived from the traditional LQG theory. This means that the LQG formulation can be handled easily in the general synthesis procedure presented in this chapter.

The next example is a H_∞-optimal control problem.

Example 2:

Figure 2-10 shows a familiar feedback control formulation. The external inputs, commands and disturbances at inputs of the plant, are assumed to be L_2-bounded signals. The performance is measured in terms of the energy of weighted tracking errors (sensitivity minimization) and weighted control input signals (preventing actuator saturation).
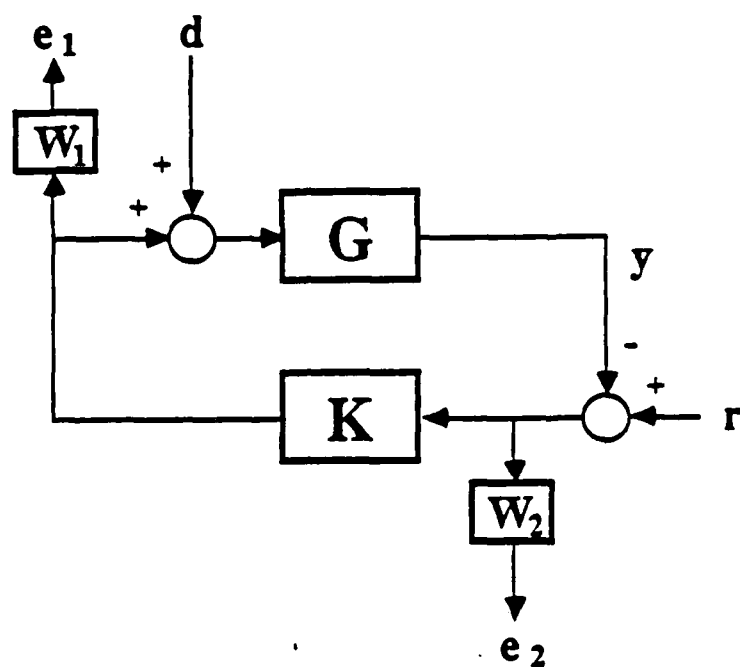
Figure 2-10. A Feedback Control Problem

Again, Figure 2-10 can be rearranged and put in the general framework as in Figure 2-11 where the interconnection matrix P is

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 & W_1 \\ -W_2 G & W_2 & -W_2 G \\ -G & I & -G \end{bmatrix}$$
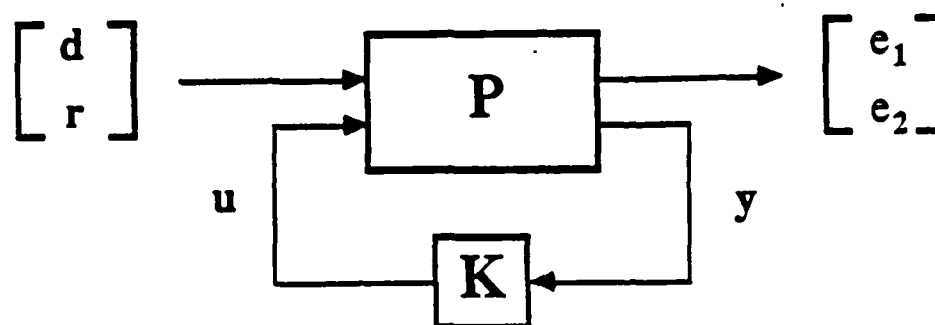
Figure 2-11.

By assumptions on the input signals and the performance, it is known [D16,F4,F5,Z2,Z3] that this is an H_∞ optimal control problem and the general distance problem arises (Eq. 2-52).

An interesting fact is that the matrix R does not necessarily have the 2×2 block structure as in Eq. 2-53. For some control problems, only two blocks appear in the GDP, that is, either $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix} = 0$ or $\begin{bmatrix} R_{21} & R_{22} \end{bmatrix} = 0$. This can be seen clearly from this example.

Note that,

$$P_{11} = \begin{bmatrix} 0 & 0 \\ -W_2 G & W_2 \end{bmatrix} = \begin{bmatrix} 0 \\ W_2 \end{bmatrix} \begin{bmatrix} -G & I \end{bmatrix} = \begin{bmatrix} 0 \\ W_2 \end{bmatrix} P_{21}$$

Then,

$$T_{11} = P_{11} + P_{12} U_* \tilde{M} P_{21} \qquad \text{(from Eq. 2-24a)}$$

$$= \left\{ \begin{bmatrix} 0 \\ W_2 \end{bmatrix} + P_{12} U_* \tilde{M} \right\} P_{21}$$

$$= \left\{ \begin{bmatrix} 0 \\ W_2 \end{bmatrix} \tilde{M}^{-1} + P_{12} U_* \right\} T_{21} \qquad \text{(from Eq. 2-24c)}$$

$$= \tilde{T}_{11} N_{21}$$

where

$$N_{21} \triangleq \bar{R}_D^{-1/2} T_{21}$$

and

$$\tilde{T}_{11} = \left\{ \begin{bmatrix} 0 \\ W_2 \end{bmatrix} \tilde{M}^{-1} + P_{12} U_* \right\} \bar{R}_D^{1/2} .$$

From Eq. 2-53,

$$R = \begin{bmatrix} N_{12}^* \\ N_1^* \end{bmatrix} T_{11} \begin{bmatrix} N_{21}^* & \bar{N}_1^* \end{bmatrix} .$$

$$= \begin{bmatrix} N_{12}^* \\ N_1^* \end{bmatrix} (\tilde{T}_{11} N_{21}) \begin{bmatrix} N_{21}^* & \bar{N}_1^* \end{bmatrix} .$$

$$= \begin{bmatrix} N_{12}^* \tilde{T}_{11} & 0 \\ N_1^* \tilde{T}_{11} & 0 \end{bmatrix} .$$

Therefore,

$$\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix} = 0.$$

This is a special general distance problem and will be referred as "2-block **GDP**. Both the 2-block and the full 2×2 block cases will be discussed in later chapters.

CHAPTER 3
GENERAL DISTANCE PROBLEMS

This chapter is devoted to a summary of results for the "General Distance Problem" (**GDP**):

Given $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \in L_\infty$, find an element $Q \in H_\infty$ such that

$$\left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty$$

is minimized. Note that the minimum value, denoted as γ_s, is the distance from R to the set of (matrix) functions of the form

$$\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad Q \in H_\infty .$$

This formulation of the **GDP** is called the "4-block problem" to distinguish it from the special cases where $\begin{bmatrix} R_{21} & R_{22} \end{bmatrix}$ or $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}$ is identically zero. This special case will be referred as the "2-block problem". Note that if both $\begin{bmatrix} R_{21} & R_{22} \end{bmatrix}$ and $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}$ are zero, this formulation is known as the "best (or Hankel) approximation" problem [A1,A2,B1,G2].

The **GDP** can also be regarded as a matrix dilation problem with the constraint of causality on Q. The subject of constant dilation will be reviewed first in Section 3.1 where Parrott's theorem on norm-preserving dilations [D1,P1,P4] will play a central role in establishing the existence of the optimal solution of the **GDP**. The *Hankel* operator and its relation in best approximation is then treated in Section 3.2 using Parrott's theorem. Section 3.3 introduces the *Toeplitz* operator and some of its algebraic properties which provide insight into the **GDP**. Generalizing the approach in Section 3.2, an abstract operator point of view is adopted. The existence of an optimal Q and expressions for the optimal norm are detailed in Section 3.4. These expressions are given in terms of an operator norm or an equivalent standard eigenvalue problem. The operator and eigenvalue problem are of infinite rank and the results do not

yield computable formulas for either the optimal norm or Q. Nevertheless, the existence of the optimal solution is established. A more practical approach will be discussed in chapter 4.

The results of this section and the next are more easily obtained using functions on the unit disc instead of the half-plane. Since there is a well-known isometric isomorphism between H_2 and H_∞ on the half-plane and the disk (see Appendix B), the general distance problem on the disk will be considered throughout this and the next chapter.

3.1 Dilation of Constant Matrix

Consider the optimization problem of finding

$$\gamma_\bullet \;=\; \min_X \;\left|\left|\begin{bmatrix} X & B \\ C & A \end{bmatrix}\right|\right|_\infty \tag{3-1}$$

where X, B, C, and A are constant matrices of compatible dimensions. The matrix $\begin{bmatrix} X & B \\ C & A \end{bmatrix}$ is a *dila-tion* of its submatrices as indicated in the following diagram:

$$
\begin{array}{ccc}
\begin{bmatrix} X & B \\ C & A \end{bmatrix} &
\begin{array}{c} c \\ \rightarrow \\ \leftarrow \\ d \end{array} &
\begin{bmatrix} B \\ A \end{bmatrix} \\[2em]
d\!\uparrow\;\downarrow\!c & & d\!\uparrow\;\downarrow\!c \\[2em]
\begin{bmatrix} C & A \end{bmatrix} &
\begin{array}{c} c \\ \rightarrow \\ \leftarrow \\ d \end{array} &
\begin{bmatrix} A \end{bmatrix}
\end{array}
\tag{3-2}
$$

Here c stands for the operation of *compression* and d stands for the operation of *dilation* . Compression is always norm decreasing; however, dilation can sometimes be made to be norm preserving. Norm preserving dilations are the focus of this section. Since it is not the purpose of this work to review the related theory throughly, only useful facts will be presented; more complete treatment on this subject can be found elsewhere [D1,D16,P1,P4].

The simplest matrix dilation problem occurs when finding

$$\gamma_\bullet \;=\; \min_X \;\left|\left|\begin{bmatrix} X \\ A \end{bmatrix}\right|\right|_\infty \tag{3-3}$$

Although Eq. 3-3 is a much simplified version of Eq. 3-1, it contains all the essential features. It is immediate that $\gamma_\bullet = ||\,A\,||_\infty$. The following theorem characterizes all solutions X to Eq. 3-3.

Theorem 1

$$\forall \gamma \geq \gamma_o ,$$

$$\left\| \left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\| \right\|_{\infty} \leq \gamma$$

if and only if

$\exists\ Y$ with $\| Y \|_{\infty} \leq 1$ (i.e., Y is a contraction) such that $X = Y(\gamma^2 I - A^*A)^{1/2}$.

This theorem implies that, in general, Eq. 3-3 has more than one solution. $X = 0$ is the central solution but others are possible unless $A^*A = \gamma^2 I$. A more restricted version of the theorem is the following corollary.

Corollary 2

$$\forall\ \gamma > \gamma_o ,$$

$$\left\| \left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\| \right\|_{\infty} \leq \gamma \quad \Longleftrightarrow \quad \| X(\gamma^2 I - A^*A)^{-1/2} \|_{\infty} \leq 1 . \tag{3-4}$$

The corresponding dual results are stated as the following theorem.

Theorem 3

$$\forall\ \gamma \geq \gamma_o ,$$

$$\| [X\ A] \|_{\infty} \leq \gamma$$

if and only if

$\exists\ Y$ with $\| Y \|_{\infty} \leq 1$ (i.e., Y is a contraction) such that $X = (\gamma^2 I - AA^*)^{1/2} Y$.

Corollary 4

$$\forall\ \gamma > \gamma_o ,$$

$$\| \, [\, X \quad A \,] \, \|_\infty \leq \gamma \quad \Longleftrightarrow \quad \| \, (\gamma^2 I - AA^*)^{-1/2} X \, \|_\infty \leq 1 . \tag{3-5}$$

The following theorem, usually attributed to Parrott [Par], characterizes γ_o and will play a central role in establishing the existence of the optimal solution to the best approximation and general distance problem.

Theorem 5

$$\gamma_o = \max \left\{ \| \, [\, C \quad A \,] \, \|_\infty , \quad \left\| \begin{bmatrix} B \\ A \end{bmatrix} \right\|_\infty \right\} \tag{3-6}$$

As in Eq. 3-3, there may be more than one solution to Eq. 3-1. The following theorem parametrizes all solutions to Eq. 3-1.

Theorem 6

Suppose $\gamma \geq \gamma_o$. The solutions X such that

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \leq \gamma \tag{3-7}$$

are exactly those of the form

$$X = -YA^*Z + \gamma(I - YY^*)^{1/2} W (I - Z^*Z)^{1/2} \tag{3-8}$$

where W is an arbitrary contraction ($\| \, W \, \|_\infty \leq 1$) and Y and Z solve the linear equations

$$B = Y(\gamma^2 I - A^*A)^{1/2} \tag{3-9}$$

and

$$C = (\gamma^2 I - AA^*)^{1/2} Z. \tag{3-10}$$

The following corollary gives an alternative version of Theorem 6.

Corollary 7

$\forall \quad \gamma > \gamma_o$,

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \leq \gamma \qquad (3\text{-}11)$$

if and only if

$$\| (I-YY^*)^{-1/2}(X + Y^*AZ)(I-Z^*Z)^{-1/2} \|_\infty \leq \gamma \qquad (3\text{-}12)$$

where

$$Y = B(\gamma^2 I - A^*A)^{-1/2}$$

$$Z = (\gamma^2 I - AA^*)^{-1/2}C. \qquad (3\text{-}13)$$

There are many alternative characterizations of solutions to Eq. 3-11, although the formula in Eq. 3-12 seems to be the simplest.

The restriction that $\gamma > \gamma_o$ in Corollary 7 does introduce some loss of generality. If these alternatives to Theorem 6 are used, it is not possible to get all solutions for $\gamma = \gamma_o$. The set of all solutions arbitrarily close to the optimal γ_o is the best that can be done. The reason for considering this special case is that Eq. 3-12 has a reasonably straightforward generalizations to the rational matrix case, whereas Eq. 3-8 does not. A further difficulty with the rational case is that, unlike the constant case, it appears that it is not possible to actually compute γ_o exactly. This makes our inability to obtain a direct generalization of Theorem 6 seem less critical, at least with respect to application of this theory.

3.2 Hankel Operators and Best Approximation Problems

Two classes of operators on a Hilbert space, *Hankel* operators and *Toeplitz* operators, have played an important role in function theory on the unit circle. In recent years the theory of *Hankel* operators has attracted increasing attention in the areas of control and systems theory, mainly due to its applications in model-reduction and best approximation.

The best (or *Hankel*) approximation is reviewed in this section where the *Hankel* operator plays an essential role. The problem is stated as follows:

For given (matrix) function $G \in L_\infty$, finding a function $Q \in H_\infty$ such that

$$\| G - Q \|_\infty \qquad \text{is minimized.}$$

This problem was first solved by Nehari for general scalar discrete systems [N1]. The multivariable version was later solved completely by Adamjan *et al.* [A1,A2]. Recently Glover considers a special case of the problem where the function to be approximated is real-rational, i.e., $G \in RL_\infty$. An efficient and constructive algorithm for the optimal solution was constructed in terms of a balanced realization of the system [G2]. For the real-rational case this is a most effective algorithm. A simple proof of existence of the minimizing solution Q is possible using a Parrott/Davis-Kahan-Weinberger theorem on norm-preserving dilations [D15,D16,P4]. This approach is quite elegant and will be used later in the proofs related to the **GDP**.

Assume that $G(z)$ belongs to $L_\infty(T)$ with the power series expansion

$$G(z) = \sum_{i=-\infty}^{\infty} G_i z^i.$$

The following definitions will be used throughout this chapter.

Definition : (Multiplicative or Laurent Operator)

The Multiplicative (Laurent) operator $\mathbf{M}_G$ generated by $G(z)$, is defined as

$$\mathbf{M}_G \;:\; L_2(T) \;\rightarrow\; L_2(T)$$

$$f \quad\rightarrow\quad \mathbf{M}_G f \;=\; G\,f$$

Definition : (*Hankel* Operator)

The *Hankel* operator $\mathbf{H}_G$ generated by $G(z)$, is defined as

$$\mathbf{H}_G \;:\; H_2(T) \;\rightarrow\; H_2^{\perp}(T)$$

$$f \quad\rightarrow\quad \mathbf{H}_G f \;=\; (\mathbf{P}_{H_2^{\perp}} \mathbf{M}_G) f$$

Remark

A matrix representation of $\mathbf{H}_G$ is

$$\begin{bmatrix} G_{-1} & G_{-2} & G_{-3} & \cdots \\ G_{-2} & G_{-3} & \bullet & \cdots \\ G_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \end{bmatrix}$$

which is the well-known (infinite) *Hankel* matrix.

Theorem 1 (Best Approximation)

Consider the following minimization problem:

$$\gamma_{\circ} \;\triangleq\; \min_{Q \,\in\, H_{\infty}} \; \| \, G \, - \, Q \, \|_{\infty} \tag{3-14}$$

Then,

$$\gamma_{\circ} \;=\; \| \, \mathbf{H}_G \, \|$$

and the minimum is achieved by some $Q \in H_{\infty}$.

Proof

The inequality

$$\inf\left\{\|G - Q\|_\infty : Q \in H_\infty(T)\right\} \geq \|\mathbf{H}_G\|$$

is easy to establish. Fix Q in $H_\infty(T)$, then

$$\|G - Q\|_\infty = \sup\left\{\|(G - Q)f\|_2 : f \in H_2(T), \|f\|_2 \leq 1\right\}$$

$$\geq \sup\left\{\|P_{H_2^\perp}(G - Q)f\|_2 : f \in H_2(T), \|f\|_2 \leq 1\right\}$$

$$= \sup\left\{\|P_{H_2^\perp}Gf\|_2 : f \in H_2(T), \|f\|_2 \leq 1\right\}$$

$$= \|\mathbf{H}_G\|.$$

and take the infimum over Q.

Since $f \in H_2(T)$, it can be written as the power series expansion

$$f(z) = \sum_{i=0}^{\infty} f_i z^i$$

and let

$$h(z) = \mathbf{H}_G f = \mathbf{P}_{H_2^\perp}(Gf) = \sum_{i=-\infty}^{-1} h_i z^i.$$

Then the matrix representation of $\mathbf{H}_G$ can be expressed as

$$\begin{bmatrix} h_{-1} \\ h_{-2} \\ h_{-3} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} G_{-1} & G_{-2} & \cdots \\ G_{-2} & G_{-3} & \cdots \\ G_{-3} & G_{-4} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}.$$

This is exactly the matrix representation of the *Hankel* operator generated by G.

Now, it is sufficient to show there exists a function Q in $H_\infty(T)$ such that

$$\| G - Q \|_\infty = \| \mathbf{H}_G \| .$$

By a result of Parrott's [P1,P4], there exists a Q_0 such that the norm of the *Hankel* matrix is preserved, i.e.,

$$\| \mathbf{H}_G \| = \left\| \left\| \begin{bmatrix} G_o -Q_o & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \right\| \right\| .$$

Similarly, there exists a Q_1 such that

$$\left\| \left\| \begin{bmatrix} G_o -Q_o & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \right\| \right\| = \left\| \left\| \begin{bmatrix} G_1 -Q_1 & G_o -Q_o & \cdots \\ G_o -Q_o & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \right\| \right\| .$$

Continuing in this way gives a suitable $Q(z) = \sum_{i=0}^{\infty} Q_i z^i \in H_\infty(T)$ such that the minimal norm is achieved.

<div align="right">QED</div>

The following corollary is immediate.

Corollary 2

$$\| \mathbf{H}_G \| = \lambda_{max}^{1/2}[\mathbf{H}_G {}^*\mathbf{H}_G] \tag{3-15}$$

where $\lambda_{max}[\mathbf{H}_G {}^*\mathbf{H}_G]$ is the maximum eigenvalue of the matrix $[\mathbf{H}_G {}^*\mathbf{H}_G]$.

Note that if G is real-rational, then there exists an optimal solution $Q \in RH_\infty$. Related materials are presented in Chapter 5 where an algorithm for constructing the optimal Q is given.

3.3 Toeplitz Operators

For the **GDP**, the solution is more complicated and both *Hankel* and *Toeplitz* operators are involved. Therefore, a overview of the necessary properties of *Toeplitz* operators will be established using simple algebraic procedures. More details on *Toeplitz* operators can be found in standard publications [B5,D7,D8,S3].

The definition of *Toeplitz* operator is given in the following.

Definition: (*Toeplitz* Operator)

The *Toeplitz* operator $\mathbf{T}_G$ generated by $G(z) \in L_\infty(T)$, is defined as

$$\mathbf{T}_G \; : \; H_2(T) \;\; \rightarrow \;\; H_2(T)$$

$$f \;\;\; \rightarrow \;\;\; \mathbf{T}_G f \;\; = \;\; (\mathbf{P}_{H_2} \mathbf{M}_G) f \;\; .$$

Remark

A matrix representation of $\mathbf{T}_G$ is

$$\begin{bmatrix} G_0 & G_1 & G_2 & \dots \\ G_{-1} & G_0 & G_1 & \dots \\ G_{-2} & G_{-1} & G_0 & \dots \\ \bullet & \bullet & \bullet & \dots \\ \bullet & \bullet & \bullet & \dots \\ \bullet & \bullet & \bullet & \dots \end{bmatrix}$$

which is the well-known (infinite) *Toeplitz* matrix.

In the following, some of algebraic properties of *Toeplitz* operators are reviewed. The first lemma is well-known.

Lemma 1

$$\| T_G \| = \| M_G \| = \| G \|_\infty$$

Proof

The proof can be found in [B5,D7].

Lemma 2

$$T_G{}^* = T_{G_*}.$$

Proof

$$\forall \ f , g \ \in H_2(T),$$

$$< T_{G_*} f , g) = (P_{H_2} M_{G_*} f , g >$$

$$= < f , M_G P_{H_2} g >$$

$$= < f , M_G g >$$

$$= < P_{H_2} f , M_G g >$$

$$= < f , P_{H_2} M_G g >$$

$$= < f , T_G g >$$

$$= < T_G{}^* f , g >$$

$$\therefore \ \ T_{G_*} = T_G{}^*.$$

QED

Definition:

The operator S_G generated by $G(z) \in L_\infty(T)$ is defined by

$$S_G \ : \ H_2^\perp(T) \ \rightarrow \ H_2(T)$$

$$f \ \rightarrow \ S_G f \ = (P_{H_2} M_G) f$$

Lemma 3

$$S_G = H_{G*}^* .$$

Proof

$\forall \ f \in H_2^\perp(T)$ and $g \in H_2(T)$,

$$< S_G f , g > \ = \ <P_{H_2}M_G f , g >$$
$$= \ < f , M_G^* P_{H_2} g >$$
$$= \ < f , M_{G*} g >$$
$$= \ < P_{H_2^\perp} f , M_{G*} g >$$
$$= \ < f , P_{H_2^\perp} M_{G*} g >$$
$$= \ < f , H_{G*} g >$$
$$= \ < H_{G*}^* f , g >$$
$$\therefore \quad S_G = H_{G*}^* .$$

QED

Lemma 4

Assume $\Phi, \Psi \in L_\infty(T)$. Then

(i) $T_{\Phi\Psi} - T_\Phi T_\Psi = H_{\Phi*}^* H_\Psi$. (Sarason)

(ii) $T_{\Phi\Psi} = T_\Phi T_\Psi$ if either Φ^* or Ψ belongs to $H_\infty(T)$. (Brown-Halmos)

Proof

(i) $\forall \ f \in H_2(T)$,

$$T_{\Phi\Psi} f \ = \ P_{H_2} M_{\Phi\Psi} f$$
$$= \ P_{H_2} M_\Phi M_\Psi f$$

$$= P_{H_2} M_\Phi \{ P_{H_2} + P_{H_2^\perp} \} M_\Psi f$$

$$= (P_{H_2} M_\Phi P_{H_2} M_\Psi) f + (P_{H_2} M_\Phi P_{H_2^\perp} M_\Psi f)$$

$$= (P_{H_2} M_\Phi P_{H_2})(P_{H_2} M_\Psi) f + (P_{H_2} M_\Phi P_{H_2^\perp})(P_{H_2} M_\Psi) f$$

$$= (T_\Phi T_\Psi + S_\Phi H_\Psi) f$$

From Lemma 3, $S_\Phi = H_{\Phi^*}^*$ and

$$T_{\Phi\Psi} = T_\Phi T_\Psi + H_{\Phi^*}^* H_\Psi$$

(ii) If either Φ^* or Ψ belongs to $H_\infty(T)$, then

$$H_{\Phi^*} = 0 \text{ or } H_\Psi = 0.$$

From (i), The equality follows immediately.

QED.

Remark

Lemma 4 shows that the *Toeplitz* operator does not have the multiplicative property in general except for the case of Lemma 4-(ii).

Definition:

The operator $\bar{T}_G$ generated by $G(z) \in L_\infty(T)$ is defined as

$$\bar{T}_G : H_2^\perp(T) \rightarrow H_2^\perp(T)$$
$$f \rightarrow \bar{T}_G f = (P_{H_2^\perp} M_G) f$$

Remark

The operator $\bar{\mathbf{T}}_G$ is unitarily equivalent to the *Toeplitz* operator $\mathbf{T}_{G_*}$ by means of the unitary operator $\mathbf{U}$ which is defined as

$$\mathbf{U} : H_2^{\perp}(T) \rightarrow H_2(T)$$

$$\mathbf{U}f = zf^*$$

that is, $\mathbf{T}_{G_*}\mathbf{U} = \mathbf{U}\,\bar{\mathbf{T}}_G$ [D9].

Lemma 5

The multiplicative (Laurent) operator $\mathbf{M}_G$ can be represented as

$$\mathbf{M}_G = \begin{bmatrix} \bar{\mathbf{T}}_G & \mathbf{H}_G \\ \mathbf{S}_G & \mathbf{T}_G \end{bmatrix}$$

Proof

Let $L_2 = H_2^{\perp} \oplus H_2$, then

$$\mathbf{M}_G : H_2^{\perp} \oplus H_2 \rightarrow H_2^{\perp} \oplus H_2$$

$$\mathbf{M}_G \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{T}}_G & \mathbf{H}_G \\ \mathbf{S}_G & \mathbf{T}_G \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

where $f_1 \in H_2^{\perp}$ and $f_2 \in H_2$.

<div align="right">QED</div>

3.4 Optimal Solutions of General Distance Problems

The fundamental issue in the general distance problem is the existence of the optimal solution. In this section it will be shown that the optimal solution of the GDP exists. Since internal stability (or equivalently, causality) is required , the nature of the problem considered will be very different from that in the constant matrix case [D1,D16]. The proof is essentially a generalization of that for best approximation given by Doyle [D15,D16] where the Parrott/Davis-Kahan-Weinberger theorem of Section 3.1 on norm-preserving dilations is used. In case that R is real-rational, it can be shown that there exists a real-rational optimal solution. The 2-block GDP will be studied first. The results are then generalized to the 4-block GDP. Both the 2 and 4 block problem have been studied by Feintuch and Prancis ([F1],[F2]). Their results are more general than those presented here because they consider time-varying as well as time-invariant systems.

Consider the following 2-block GDP:

$$\gamma_o = \min_{Q \in H_\infty} \left\| R - \begin{bmatrix} Q \\ 0 \end{bmatrix} \right\|_\infty \quad \text{where} \quad R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \in L_\infty. \tag{2gdp}$$

Define the operator Γ_R from H_2 to $H_2^\perp \oplus L_2$ as follows:

$$\Gamma_R : H_2 \rightarrow H_2^\perp \oplus L_2$$

$$\Gamma_R f = \begin{bmatrix} P_{H_2^\perp} R_1 f \\ R_2 f \end{bmatrix} , \quad f \in H_2 \tag{3-16}$$

Theorem 1

Consider the 2-block GDP in (2gdp), then

$$\min_{Q \in H_\infty} \left\| \begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix} \right\|_\infty = \| \Gamma_R \| .$$

Proof

The following inequality is proved first.

$$\min_{Q \in H_\infty} \left|\left|\begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix}\right|\right|_\infty \geq ||\Gamma_R|| \tag{3-17}$$

$\forall \; Q \in H_\infty$,

$$\left|\left|\begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix}\right|\right|_\infty = \sup\left\{ \left|\left|\begin{bmatrix} (R_1 - Q)f \\ R_2 f \end{bmatrix}\right|\right|_2 : f \in H_2, \; ||f||_2 \leq 1 \right\}$$

$$\geq \sup\left\{ \left|\left|\begin{bmatrix} \mathbf{P}_{H_2^\perp}(R_1 - Q)f \\ R_2 f \end{bmatrix}\right|\right|_2 : f \in H_2, \; ||f||_2 \leq 1 \right\}$$

$$= \sup\left\{ \left|\left|\begin{bmatrix} \mathbf{P}_{H_2^\perp} R_1 f \\ R_2 f \end{bmatrix}\right|\right|_2 : f \in H_2, \; ||f||_2 \leq 1 \right\}$$

$$= ||\Gamma_R||$$

It will be convenient to complete the proof by transforming the *rhp* into the unit disc (see Appendix B). For the rest of the proof assume that this has been done. Since $R_1, R_2 \in L_\infty(T)$, they can be written as power series expansions:

$$R_1(z) = \sum_{t = -\infty}^{\infty} \alpha_t z^t$$

$$R_2(z) = \sum_{t = -\infty}^{\infty} \beta_t z^t$$

and

$$f(z) = \sum_{i = 0}^{\infty} f_i z^i \quad \in H_2$$

Now, let

$$h_1(z) = \sum_{-\infty}^{-1} (h_1)_i z^i$$

$$= (R_1 - Q)f \Big|_{\mathbf{P}_{H_2^\perp}} = (R_1 f)\Big|_{\mathbf{P}_{H_2^\perp}} = \left(\sum_{t = -\infty}^{\infty} \alpha_t z^t \sum_{i = 0}^{\infty} f_i z^i \right) \Big|_{\mathbf{P}_{H_2^\perp}}$$

$$h_2(z) = \sum_{i=-\infty}^{\infty} (h_2)_i z^i$$

$$= R_2(z) f(z) = \sum_{k=-\infty}^{\infty} \beta_k z^k \sum_{i=0}^{\infty} f_i z^i$$

Using these expression for $h_1(z)$ and $h_2(z)$, Eq. 3-16 can be written as the following equivalent matrix equation:

$$
\begin{bmatrix}
(h_1)_{-1} \\
(h_1)_{-2} \\
(h_1)_{-3} \\
\bullet \\
\bullet \\
\bullet \\
(h_2)_2 \\
(h_2)_1 \\
(h_2)_0 \\
(h_2)_{-1} \\
(h_2)_{-2} \\
(h_2)_{-3} \\
\bullet \\
\bullet
\end{bmatrix}
=
\begin{bmatrix}
\alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\
\alpha_{-2} & \alpha_{-3} & \bullet & \cdots \\
\alpha_{-3} & \bullet & \bullet & \cdots \\
\bullet & \bullet & \bullet & \cdots \\
\bullet & \bullet & \bullet & \cdots \\
\bullet & \bullet & \beta_1 & \cdots \\
\bullet & \beta_1 & \beta_0 & \cdots \\
\beta_1 & \beta_0 & \beta_{-1} & \cdots \\
\beta_0 & \beta_{-1} & \beta_{-2} & \cdots \\
\beta_{-1} & \beta_{-2} & \beta_{-3} & \cdots \\
\beta_{-2} & \beta_{-3} & \bullet & \cdots \\
\beta_{-3} & \bullet & \bullet & \cdots \\
\bullet & \bullet & \bullet & \cdots \\
\bullet & \bullet & \bullet & \cdots
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1 \\
f_2 \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet
\end{bmatrix}
$$

The next step is to show that there exists an optimal $Q \in H_\infty$ such that the equality in Eq. 3-17 holds; for which
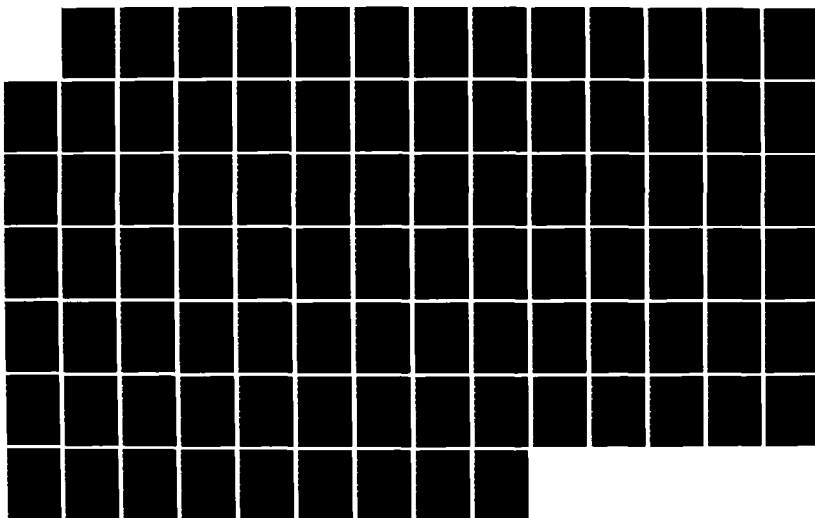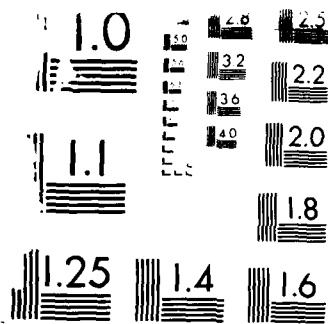
$$\|\Gamma_R\| = \left\| \begin{bmatrix} \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_{-2} & \alpha_{-3} & \bullet & \cdots \\ \alpha_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \beta_1 & \cdots \\ \bullet & \beta_1 & \beta_0 & \cdots \\ \beta_1 & \beta_0 & \beta_{-1} & \cdots \\ \beta_0 & \beta_{-1} & \beta_{-2} & \cdots \\ \beta_{-1} & \beta_{-2} & \beta_{-3} & \cdots \\ \beta_{-2} & \beta_{-3} & \bullet & \cdots \\ \beta_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \end{bmatrix} \right\|_\infty \tag{3-18}$$

From Theorem 3.1.5 (Parrott's theorem), there exists a Q_0 such that

$$\|\Gamma_R\| = \left\| \begin{bmatrix} \alpha_0 - Q_0 & \alpha_{-1} & \alpha_{-2} & \cdots \\ \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_{-2} & \alpha_{-3} & \bullet & \cdots \\ \alpha_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \beta_1 & \cdots \\ \bullet & \beta_1 & \beta_0 & \cdots \\ \beta_1 & \beta_0 & \beta_{-1} & \cdots \\ \beta_0 & \beta_{-1} & \beta_{-2} & \cdots \\ \beta_{-1} & \beta_{-2} & \beta_{-3} & \cdots \\ \beta_{-2} & \beta_{-3} & \bullet & \cdots \\ \beta_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \end{bmatrix} \right\|_\infty$$

Next, choose Q_1 such that

$$||\Gamma_R|| = \left\Vert \left[\begin{array}{cccc} \alpha_1 - Q_1 & \alpha_0 - Q_0 & \alpha_{-1} & \cdots \\ \alpha_0 - Q_0 & \alpha_{-1} & \alpha_{-2} & \cdots \\ \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_{-2} & \alpha_{-3} & \bullet & \cdots \\ \alpha_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \beta_1 & \cdots \\ \bullet & \beta_1 & \beta_0 & \cdots \\ \beta_1 & \beta_0 & \beta_{-1} & \cdots \\ \beta_0 & \beta_{-1} & \beta_{-2} & \cdots \\ \beta_{-1} & \beta_{-2} & \beta_{-3} & \cdots \\ \beta_{-2} & \beta_{-3} & \bullet & \cdots \\ \beta_{-3} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \end{array} \right] \right\Vert_\infty$$

Continuing in this way to find $\{Q_0, Q_1, Q_2, \ldots\ldots\}$ such that

$$\gamma_o = \left\Vert \left[\begin{array}{c} R_1 - Q_{opt} \\ R_2 \end{array} \right] \right\Vert_\infty = ||\Gamma_R||, \quad \text{where} \quad Q_{opt}(z) = \sum_{i=0}^{\infty} Q_i z^i \in H_\infty$$

This concludes the proof of equality.

<div align="right">QED</div>

The minimal norm γ_o can also be expressed in terms of the following eigenvalue problem ([V4]).

Corollary 2

$$||\Gamma_R|| = \lambda_{max}^{1/2}\left[H_{R_1}{}^* H_{R_1} + T_{R_2^* R_2} \right] \tag{3-19}$$

where H_{R_1} is the *Hankel* matrix generated by R_1 and $T_{R_2^* R_2}$ is the *Toeplitz* matrix generated by $R_2^* R_2$

Proof

Assume that $\hat{R}_2$ is a function in H_∞ such that $\hat{R}_2^*\hat{R}_2 = R_2^*R_2$ and define

$$\hat{R} = \begin{bmatrix} R_1 \\ \hat{R}_2 \end{bmatrix} \quad \text{where} \quad \hat{R}_2(z) = \sum_{i=0}^{\infty} \hat{\beta}_i z^i .$$

Following the proof in Theorem 1, there exists a $\hat{Q}_{opt}(z) = \sum_{i=0}^{\infty} \hat{Q}_i z^i \in H_\infty$ such that

$$\|\Gamma_{\hat{R}}\| = \min_{\hat{Q} \in H_\infty} \left\| \begin{bmatrix} R_1 - \hat{Q} \\ \hat{R}_2 \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} R_1 - \hat{Q}_{opt} \\ \hat{R}_2 \end{bmatrix} \right\|_\infty$$

Since

$$\left\| \begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} R_1 - Q \\ \hat{R}_2 \end{bmatrix} \right\|_\infty \quad \forall \ Q \in H_\infty$$

$$==> \qquad \gamma_o = \|\Gamma_{\hat{R}}\|$$

Therefore,

$$\gamma_o = \left\| \begin{bmatrix} R_1 - \hat{Q}_{opt} \\ \hat{R}_2 \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} \begin{matrix} \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \alpha_1 - \hat{Q}_1 & \alpha_0 - \hat{Q}_0 & \alpha_{-1} & \cdots \\ \alpha_0 - \hat{Q}_0 & \alpha_{-1} & \alpha_{-2} & \cdots \\ \alpha_{-1} & \alpha_{-2} & \bullet & \cdots \\ \alpha_{-2} & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \hat{\beta}_1 & \cdots \\ \bullet & \hat{\beta}_1 & \hat{\beta}_0 & \cdots \\ \hat{\beta}_1 & \hat{\beta}_0 & 0 & \cdots \\ \hat{\beta}_0 & 0 & 0 & \cdots \\ 0 & 0 & \bullet & \cdots \\ 0 & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \end{matrix} \end{bmatrix} \right\|_\infty \tag{3-20}$$

The right hand side of Eq. 3-20 is also equal to

$$\left\| \left\| \begin{bmatrix} \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_{-2} & \alpha_{-3} & \cdot & \cdots \\ \alpha_{-3} & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \cdots \\ \cdot & \cdot & \hat{\beta}_1 & \cdots \\ \cdot & \hat{\beta}_1 & \hat{\beta}_0 & \cdots \\ \hat{\beta}_1 & \hat{\beta}_0 & 0 & \cdots \\ \hat{\beta}_0 & 0 & 0 & \cdots \\ 0 & 0 & \cdot & \cdots \\ 0 & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{bmatrix} \right\| \right\|_\infty \qquad (3\text{-}21)$$

because $\{\hat{Q}_i\}$ are chosen to be norm-preserving. The zero rows in Eq. 3-21 can then be deleted without affecting the norm.

$$\therefore \qquad \gamma_o = \left\| \begin{bmatrix} H_{R_1} \\ T_{\hat{R}_2} \end{bmatrix} \right\|_\infty = \lambda_{\max}^{1/2} \left[H_{R_1}{}^* H_{R_1} + T_{\hat{R}_2}{}^* T_{\hat{R}_2} \right]$$

Since $\hat{R}_2 \in H_\infty$, by Lemmas 3.3.2 and 3.3.4-(ii),

$$T_{\hat{R}_2}{}^* T_{\hat{R}_2} = T_{\hat{R}_2{}^*} T_{\hat{R}_2} = T_{\hat{R}_2{}^* \hat{R}_2} = T_{R_2{}^* R_2}$$

This completes the proof.

QED

Remark

Corollary 2 can also be proved using Lemma 3.3.4-(i) (Sarason [S3]). Recognizing that Eq. 3-18 is equivalent to

$$\gamma_o = \left\| \begin{bmatrix} H_{R_1} \\ T_{R_2} \\ H_{R_2} \end{bmatrix} \right\|_\infty = \lambda_{\max}^{1/2} \left[H_{R_1}{}^* H_{R_1} + T_{R_2}{}^* T_{R_2} + H_{R_2}{}^* H_{R_2} \right]$$

The result follows immediately from Lemma 3.3.4-(i), since

$$T_{R_2^\sharp R_2} - T_{R_2}^* T_{R_2} = T_{R_2^\sharp R_2} - T_{R_2^\sharp} T_{R_2} = H_{(R_2^\sharp)^*}^* H_{R_2} = H_{R_2}^* H_{R_2} .$$

The following theorem is the real-rational version of Theorem 1. The theorem is stated without proof which will be given in Chapter 5 where real-rational matrices are the focused topic.

Theorem 3

If $R \in RL_\infty$ (real-rational) in (2gdp), then there exists an optimal $Q \in RH_\infty$.

Although Theorem 3.1 and Corollary 3.2 give explicit formulas for γ_*, the formulas cannot be used directly to compute either γ_* or optimal Q's. Certain observations can be made from general operator theoretic considerations which indicate the difficulty of such computations. A Hermitian *Toeplitz* operator has no point spectrum (i.e., no eigenvalues). This is known as the Hartman-Winter theorem (see Douglas ([D9])). Therefore, in Eq. 3-4, $T_{R_2^\sharp R_2}$ has infinite rank even when R_2 is rational. This is quite different from the simple best-approximation problem (e.g. Adamjan, Arov, and Krein ([A1],[A2]), and Glover ([G2])), where in the real-rational case the corresponding *Hankel* matrix has only finite rank which is equal to the McMillan degree of the given transfer matrix. While neither Theorem 3.1 nor Corollary 3.2 provide a numerically attractive method for computating γ_*, which remains an open question, existence is settled, and the *Hankel* $\oplus$ *Toeplitz* structure appearing in Corollary 3.2 can be analyzed to provide further insight into the problem of computing γ_*.

Since for rational R the Hankel part of the operator in (3-4) is finite rank and therefore compact, the operator $H_{R_1}^* H_{R_1} + T_{R_2^\sharp R_2}$ can be viewed as a compact perturbation $H_{R_1}^* H_{R_1}$ of the operator $T_{R_2^\sharp R_2}$. A standard result in operator theory is that compact perturbations do not change the continuous part of the spectrum of an operator (Gohberg and Krein, [G4]). It is easily shown from these general operator theoretic considerations that the spectrum of $H_{R_1}^* H_{R_1} + T_{R_2^\sharp R_2}$ consists of the continuous spectrum of $T_{R_2^\sharp R_2}$ plus a finite number of discrete elements. The spectral radius is then either equal to that of $T_{R_2^\sharp R_2}$ and can be computed from $\rho^{1/2}\left[T_{R_2^\sharp R_2}\right] = || R_2 ||_\infty$ or is achieved on the discrete part of

the spectrum. In the latter case, there is no known method for directly obtaining γ_o; this has motivated the investigation involving approximations and bounds that appear later in this section as well as the iterative scheme in the next section.

Operators with the *Hankel $\oplus$ Toeplitz* structure of Corollary 3.2 have been investigated in some detail by Jonckheere and co-workers, both in the H_∞ context ([V4]) and earlier in problems arising in LQG (e.g. [J1]). Recently, Jonkheere and Verma ([J2]) have proposed a scheme for estimating γ_o by computing the solution to a single Riccati equation; it is hoped that this will ultimately lead to more computationally attractive methods than are currently available.

To avoid these difficulties which arise in the direct approach, an iterative scheme, called a γ-iteration, will be proposed in Chapter 4. The rest of this section is devoted to the generalization of Theorem 1 to the 4-block **GDP**.

For the 4-block **GDP**, i.e.,

$$\min_{Q \in H_\infty} \left\| \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\| \right\|_\infty \tag{4gdp}$$

let's define the operator Γ_R as follows:

$$\Gamma_R \; : \; H_2 \oplus L_2 \; \rightarrow \; H_2^\perp \oplus L_2$$

$$\Gamma_R \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{H_2^\perp} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}.$$

Theorem 4

- Consider the 4-block **GDP** and assume

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \in L_\infty .$$

Then

$$\min_{Q \in H_\infty} \left\| \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\| \right\|_\infty = \| \Gamma_R \|$$

Proof

The proof is based on the Parrott/Davis-Kahan-Weinberger theorem, and follows closely the proof of Theorem 1.

First it is useful to establish that

$$\min_{Q \in H_\infty} \left\| \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\| \right\|_\infty \geq \| \Gamma_R \| \tag{3-22}$$

$\forall \quad Q \in H_\infty$,

$$\left\| \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\| \right\|_\infty$$

$$= \sup \left\{ \left\| \left\| \begin{bmatrix} (R_{11} - Q) f + R_{12} g \\ R_{21} f + R_{22} g \end{bmatrix} \right\| \right\|_2 : f \in H_2, g \in L_2, \left\| \left\| \begin{matrix} f \\ g \end{matrix} \right\| \right\|_2 \leq 1 \right\}$$

$$\geq \sup \left\{ \left\| \left\| \begin{bmatrix} \mathbf{P}_{H_2^\perp} \{ (R_{11} - Q) f + R_{12} g \} \\ R_{21} f + R_{22} g \end{bmatrix} \right\| \right\|_2 : f \in H_2, g \in L_2, \left\| \left\| \begin{matrix} f \\ g \end{matrix} \right\| \right\|_2 \leq 1 \right\}$$

$$= \sup \left\{ \left\| \left\| \begin{bmatrix} \mathbf{P}_{H_2^\perp} (R_{11} f + R_{12} g) \\ R_{21} f + R_{22} g \end{bmatrix} \right\| \right\|_2 : f \in H_2, g \in L_2, \left\| \left\| \begin{matrix} f \\ g \end{matrix} \right\| \right\|_2 \leq 1 \right\}$$

$$= \| \Gamma_R \| .$$

The matrix representation of the operator Γ_R is derived as follows. Since $R \in L_\infty$, its four submatrices can be written as power series expansions:

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} \sum_{k=-\infty}^{\infty} \alpha_k z^k & \sum_{k=-\infty}^{\infty} \eta_k z^k \\ \sum_{k=-\infty}^{\infty} \beta_k z^k & \sum_{k=-\infty}^{\infty} \zeta_k z^k \end{bmatrix}$$

$$f(z) = \sum_{i=0}^{\infty} f_i z^i \quad \in H_2$$

and

$$g(z) = \sum_{i=-\infty}^{\infty} g_i z^i \quad \in L_2$$

Now, let

$$h_1(z) = \sum_{i=-\infty}^{-1} (h_1)_i z^i$$

$$= (R_{11} - Q)f + R_{12}g\Big|_{P_{H_2^\perp}}$$

$$= (R_{11}f + R_{12}g)\Big|_{P_{H_2^\perp}}$$

$$= \left(\sum_{k=-\infty}^{\infty} \alpha_k z^k \sum_{i=0}^{\infty} f_i z^i + \sum_{k=-\infty}^{\infty} \eta_k z^k \sum_{i=-\infty}^{\infty} g_i z^i \right) \Big|_{P_{H_2^\perp}}$$

$$h_2(z) = \sum_{i=-\infty}^{\infty} (h_2)_i z^i$$

$$= R_{21}f + R_{22}g$$

$$= \sum_{k=-\infty}^{\infty} \beta_k z^k \sum_{i=0}^{\infty} f_i z^i + \sum_{k=-\infty}^{\infty} \zeta_k z^k \sum_{i=-\infty}^{\infty} g_i z^i$$

which can be expressed in the following matrix form:

$$
\begin{bmatrix}
(h_1)_{-1} \\
(h_1)_{-2} \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
(h_2)_1 \\
(h_2)_0 \\
(h_2)_{-1} \\
(h_2)_{-2} \\
\bullet \\
\bullet
\end{bmatrix}
=
\begin{bmatrix}
\alpha_{-1} & \alpha_{-2} & & \eta_1 & \eta_0 & \eta_{-1} & \eta_{-2} \\
\alpha_{-2} & \bullet & & \eta_0 & \eta_{-1} & \eta_{-2} & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet \\
\beta_1 & \beta_0 & & \bullet & \bullet & \zeta_1 & \zeta_0 \\
\beta_0 & \beta_{-1} & & \bullet & \zeta_1 & \zeta_0 & \zeta_{-1} \\
\beta_{-1} & \beta_{-2} & & \zeta_1 & \zeta_0 & \zeta_{-1} & \zeta_{-2} \\
\beta_{-2} & \bullet & & \zeta_0 & \zeta_{-1} & \zeta_{-2} & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & & \bullet & \bullet & \bullet & \bullet
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1 \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
g_{-2} \\
g_{-1} \\
g_0 \\
g_1 \\
\bullet \\
\bullet
\end{bmatrix}
$$

The next step is to show that there exists an optimal $Q \in H_\infty$ such that the minimum is achieved. Therefore

$$\gamma_o = \|\Gamma_R\| = \left\| \left\| \begin{bmatrix} \alpha_{-1} & \alpha_{-2} & & & \eta_1 & \eta_0 & \eta_{-1} & \eta_{-2} \\ \alpha_{-2} & \bullet & & & \eta_0 & \eta_{-1} & \eta_{-2} & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \\ \beta_1 & \beta_0 & & & \bullet & \bullet & \zeta_1 & \zeta_0 \\ \beta_0 & \beta_{-1} & & & \bullet & \zeta_1 & \zeta_0 & \zeta_{-1} \\ \beta_{-1} & \beta_{-2} & & & \zeta_1 & \zeta_0 & \zeta_{-1} & \zeta_{-2} \\ \beta_{-2} & \bullet & & & \zeta_0 & \zeta_{-1} & \zeta_{-2} & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & & \bullet & \bullet & \bullet & \bullet \end{bmatrix} \right\| \right\|_\infty \tag{3-23}$$

Applying the norm-preserving dilation repeatedly, $\{Q_0, Q_1, Q_2 \ldots\ldots\}$ can be found such that

$$\gamma_o = \left\| \left\| \begin{bmatrix} R_{11} - Q_{opt} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\| \right\|_\infty \quad \text{where} \quad Q_{opt}(z) = \sum_{i=0}^{\infty} Q_i z^i \in H_\infty .$$

This concludes the proof of equality.

QED

Identifying the structure in Eq. 3-23 with the operators defined in Section 3.2 and 3.3, the following corollary follows immediately.

Corollary 5

$$\gamma_o = \left\| \left\| \begin{bmatrix} H_{R_{11}} & \tilde{T}_{R_{12}} & H_{R_{12}} \\ T_{R_{21}} & S_{R_{22}} & T_{R_{22}} \\ H_{R_{21}} & \tilde{T}_{R_{22}} & H_{R_{22}} \end{bmatrix} \right\| \right\| \tag{3-24}$$

Similar to 2-block case (Theorem 3), the following theorem will be proved later in Chapter 5.

Theorem 6

If $R \in RL_\infty(T)$ (real-rational) in (4gdp), then there exists an optimal $Q \in RH_\infty$.

Although the existence of the optimal solution is also justified for the 4-block case, Theorem 3 suffers the same criticism as in the 2-block case of being nonconstructive. The associated eigenvalue problem (Corollary 5) is much worse than that in Corollary 2. An alternative approach must be sought. This will be the subject of the next chapter.

CHAPTER 4
AN ITERATIVE METHOD FOR
GENERAL DISTANCE PROBLEMS

The existence of the optimal solution of the general distance problem (GDP) is shown in Chapter 3 from a somewhat abstract operator theory point of view. Although the approach is conceptually elegant, it doesn't give any computable formula for either the minimal achievable norm γ_o or the optimal solution Q. In this chapter an alternative approach, γ-iteration, is introduced which involves guessing a γ and then reducing the problem of finding all Q that give norm less than γ to a standard best (Hankel-norm) approximation problem. In Section 4.1 a method for reducing the **GDP** to an equivalent best approximation problem is shown Bounds for the optimal γ_o are given in terms of easily computable quantities which yield reasonable estimates of γ_o. The guess for γ is iterated on until it converges to the optimal norm, γ_o. The optimal Q is thus obtained. A general description of the γ-iteration procedure is presented. In Section 4.2, the γ-iteration is viewed as finding the "zero crossing" of a function. This function is shown to be continuous, monotonically decreasing, convex, and bounded by some very simple functions. These properties make very rapid convergence of the γ-iteration possible. The conditions for which the optimal norm is achieved are given. An example is shown in Section 4.3 to illustrate the important aspects of the **GDP.**

4.1 γ-Iteration and Bounds

In this section, we present an iterative scheme suggested in Doyle (1983,1984) to solve the general distance problem. The idea is that by guessing a value for the minimal norm, γ_o, the distance problem can be simplified to an equivalent best approximation problem which can be solved by existing algorithms (for example, Glover (1984)). This guess can be iterated to obtain convergence to the optimal norm and optimal Q. This γ-iteration procedure was suggested independently by Francis (1983) (Chapter 8) for the 2-block GDP and by Verma and Jonckheere (1984) for the so-called "SISO mixed sensitivity problem," a 2-block case.

Again, the 2-block and 4-block cases will be considered separately. Theorems 4.1 and 4.2 lie at the heart of the γ-iteration scheme. In what follows, M will be called a spectral factor of a rational para-Hermitian S if $M \in RH_\infty$ is a unit ($M^{-1} \in RH_\infty$), and $M^*M = S$.

Theorem 4.1

Assume $Q \in H_\infty$ and $\gamma > \| R_2 \|_\infty$. Then

$$\left\| \left[\begin{matrix} R_1 - Q \\ R_2 \end{matrix} \right] \right\|_\infty \leq \gamma \tag{4-1}$$

if and only if

$$\| (R_1 - Q) M^{-1} \|_\infty \leq 1 \tag{4-2}$$

where M is a spectral factor of the para-Hermitian matrix $(\gamma^2 I - R_2^* R_2)$.

Proof

Note that

$$(R_1 - Q)^*(R_1 - Q) \leq \gamma^2 I - R_2^* R_2$$

evaluated on the imaginary axis is equivalent to both (4-1) and (4-2).

QED

The theorem means that if $\gamma > ||R_2||_\infty$, there exists a $\bar{Q} \in H_\infty$ such that $||R_1M^{-1} - \bar{Q}||_\infty \leq 1$ and hence, $||\mathbf{H}_{R_1M^{-1}}|| \leq 1$. For such a $\bar{Q}$, $\bar{Q}M$ will satisfy Eq. 4-1. Solutions to Eq. 4-2 can be obtained by considering the following best approximation problem

$$f(\gamma) = \min_{Q \in H_\infty} ||R_1M^{-1} - \bar{Q}||_\infty = ||\mathbf{H}_{R_1M^{-1}}||. \tag{4-3}$$

Therefore, the general distance problem can be reduced to a best approximation problem. In the case that the function R_1M^{-1} is real-rational, the algorithms in Glover (1984) can then be applied to solve for the optimal $Q \in RH_\infty$ corresponding to the given γ.

The case where $\gamma = ||R_2||_\infty$ must be handled slightly differently because the factorization required in Theorem 4.1 would yield an M with zeros on the $j\omega$-axis and thus not a unit. A necessary condition for 4-2 to hold is that $R_1M^{-1} \in L_\infty$ even though M^{-1} is not in L_∞. While this is possible conceptually it is a condition that can never be verified numerically and we must settle in practice for the conditions of Theorem 4.1.

Using Theorem 4.1, we can view $\bar{\gamma}$ in 4-3 as a function of γ and iterate on the choice of γ until $\bar{\gamma} = 1$. The following procedure is a general description of the γ-iteration procedure for the 2-block problem:

(i)   Compute the lower bound $||R_2||_\infty$.

(ii) Choose γ such that $||R_2||_\infty < \gamma$.

(iii) Find a spectral factor $M = (\gamma^2 I - R_2^* R_2)^{1/2}$.

(iv) Let $\bar{\gamma} = ||\mathbf{H}_{R_1M^{-1}}||$,

      (a) if $\bar{\gamma} > 1$, go to (ii) and choose a larger γ.

      (b) if $\bar{\gamma} < 1$, go to (ii) and choose a smaller γ.

(c) if $\bar{\gamma}=1$ go to (v).

(v)   The value of γ is the minimal achievable norm. Find a best approximation of $R_1 M^{-1}$, denoted by $\tilde{Q}_o$.

(vi) An optimal solution $Q_{opt} = \tilde{Q}_{opt} M$.

A numerical implementation of this procedure could not, of course, evaluate $\bar{\gamma}$ exactly, so the test in (iv-c) would never be met exactly. This procedure could be used to find a Q that yields a solution arbitrarily close to γ_o. In order to do this, some scheme for selecting the next guess for γ in step (iv) is needed and the convergence properties of the procedure would depend on this scheme. The next section will focus on the properties of $\bar{\gamma}$ as a function of γ that make it possible to converge rapidly to γ_o.

Since the approach proposed is an iterative one, it will be helpful if the upper and lower bounds can be provided in advance. The bounds in Theorems 2 and 4 are quite useful since they are relatively easily computed and the upper and lower bounds have a ratio no greater than $\sqrt{2}$ in the 2-block case (Theorem 2) and 2 in the 4-block case (Theorem 4).

Theorem 2

Let

$$\gamma_o = \min_{Q \in H_\infty} \left\| \left\| \begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix} \right\| \right\|_\infty$$

$$\gamma_u = \left\| \left\| \begin{bmatrix} \|\mathbf{H}_{R_1}\| \\ \|R_2\|_\infty \end{bmatrix} \right\| \right\|_2$$

$$\gamma_{l_1} = \max \left\{ \|\mathbf{H}_R\|, \|R_2\|_\infty \right\}$$

$$\gamma_{l_2} = \max \left\{ \|\mathbf{H}_{R_1}\|, \|R_2\|_\infty \right\}$$

Then

$$\frac{1}{\sqrt{2}}\,\gamma_o \leq \gamma_{l_2} \leq \gamma_{l_1} \leq \gamma_o \leq \gamma_u \leq \sqrt{2}\,\gamma_o$$

Proof

Let Q_{opt} and $\tilde{Q}_{opt}$ satisfy

$$\left\|\begin{bmatrix} R_1 - Q_{opt} \\ R_2 \end{bmatrix}\right\|_\infty = \min_{Q \in H_\infty} \left\|\begin{bmatrix} R_1 - Q \\ R_2 \end{bmatrix}\right\|_\infty$$

and

$$\| R_1 - \tilde{Q}_{opt} \|_\infty = \min_{Q \in H_\infty} \| R_1 - \tilde{Q} \|_\infty$$

To prove that $\gamma_o \leq \gamma_u \leq \sqrt{2}\,\gamma_o$, note that

$$\left\|\begin{bmatrix} R_1 - Q_{opt} \\ R_2 \end{bmatrix}\right\|_\infty \leq \left\|\begin{bmatrix} R_1 - \tilde{Q}_{opt} \\ R_2 \end{bmatrix}\right\|_\infty$$

$$\leq \left\|\begin{bmatrix} \| R_1 - \tilde{Q}_{opt} \|_\infty \\ \|R_2\|_\infty \end{bmatrix}\right\|_2 \quad (=\gamma_u) \qquad \text{(from Lemma A.1)}$$

$$\leq \left\|\begin{bmatrix} \| R_1 - Q_{opt} \|_\infty \\ \|R_2\|_\infty \end{bmatrix}\right\|_2$$

$$\leq \sqrt{2}\left\|\begin{bmatrix} R_1 - Q_{opt} \\ R_2 \end{bmatrix}\right\|_\infty \qquad \text{(from Lemma A.1)}$$

$$\therefore \quad \gamma_o \leq \gamma_u \leq \sqrt{2}\,\gamma_o$$

To prove that $\gamma_{l_2} \leq \gamma_{l_1} \leq \gamma_o$, note that $\|R_2\|_\infty \leq \gamma_o$ and

$$\|\mathbf{H}_{R_1}\| \leq \|\mathbf{H}_R\| = \min_{Q \in H_\infty} \| R - \tilde{Q} \|_\infty \leq \left\|R - \begin{bmatrix} Q_o \\ 0 \end{bmatrix}\right\|_\infty = \gamma_o$$

Finally to prove that $\frac{1}{\sqrt{2}}\,\gamma_o \leq \gamma_{l_2}$, note that

$$\gamma_o \leq \gamma_u = \left\|\begin{bmatrix} \|\mathbf{H}_{R_1}\| \\ \|R_2\|_\infty \end{bmatrix}\right\|_2 \leq \sqrt{2}\max\left\{\|\mathbf{H}_{R_1}\|\,,\,\|R_2\|_\infty\right\} = \gamma_{l_2}$$

$$\therefore \quad \frac{1}{\sqrt{2}}\gamma_o \leq \gamma_{l_2}.$$

<div align="right">**QED**</div>

The following two theorems are the generalizations of Theorems 1 and 2 to the 4-block **GDP**.

Theorem 3

$$\left|\left|\begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\right|\right|_\infty \leq \gamma \qquad (\gamma > \gamma_{l_1}) \tag{4-6}$$

if and only if

$$\left|\left|(I - LL^*)^{-1/2}\left\{ \mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^*) - \frac{1}{\gamma}Q \right\}(I - \tilde{L}^*\tilde{L})^{-1/2}\right|\right|_\infty \leq 1 \tag{4-7}$$

where

$$(I - LL^*)^{1/2} = \text{spectral factor of} \ (I - LL^*)$$

$$(I - \tilde{L}^*\tilde{L})^{1/2} = \text{spectral factor of} \ (I - \tilde{L}^*\tilde{L})$$

$$L = R_{12}S^{-1}$$

$$\tilde{L} = \tilde{S}^{-1}R_{21}$$

$$S = (\gamma^2 I - R_{22}^* R_{22})^{1/2}$$

$$\tilde{S} = (\gamma^2 I - R_{22}R_{22}^*)^{1/2}$$

and

$$\mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^*) = \frac{1}{\gamma}\left\{ R_{11} + R_{12}(\gamma^2 I - R_{22}^* R_{22})^{-1}R_{22}^* R_{21}\right\}$$

Proof

From Theorem 1,

$$\left|\left|\begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\right|\right|_\infty \leq \gamma \qquad (\gamma > \gamma_o)$$

$\longleftrightarrow$

$$\left|\left|\left[R_{11} - Q \ \ R_{12}\right] T^{-1}\right|\right|_{\infty} \leq 1 \tag{4-8}$$

where T satisfies $T^{*}T = (\gamma^2 I - \begin{bmatrix} R_{21}^{*} \\ R_{22}^{*} \end{bmatrix} \begin{bmatrix} R_{21} & R_{22} \end{bmatrix})$ and can be chosen, from Lemma A.3, as

$$T = \begin{bmatrix} \gamma(I - \tilde{L}^{*}\tilde{L})^{1/2} & 0 \\ -(S^{-1})^{*}R_{22}^{*}R_{21} & S \end{bmatrix}$$

$$\therefore \quad \left[R_{11} - Q \ \ R_{12}\right] T^{-1}$$

$$= \left[R_{11} - Q \ \ R_{12}\right] \begin{bmatrix} \dfrac{1}{\gamma}(I - \tilde{L}^{*}\tilde{L})^{-1/2} & 0 \\ \dfrac{1}{\gamma}(\gamma^2 I - R_{22}^{*}R_{22})^{-1}R_{22}^{*}R_{21}(I - \tilde{L}^{*}\tilde{L})^{-1/2} & S^{-1} \end{bmatrix}$$

$$= \left[\dfrac{1}{\gamma}\left\{ \left(R_{11} + R_{12}(\gamma^2 I - R_{22}^{*}R_{22})^{-1}R_{22}^{*}R_{21}\right) - Q \right\}(I - \tilde{L}^{*}\tilde{L})^{-1/2} \ \ R_{12}S^{-1} \right]$$

Recognizing that

$$\frac{1}{\gamma}\left(R_{11} + R_{12}(\gamma^2 I - R_{22}^{*}R_{22})^{-1}R_{22}^{*}R_{21}\right) = \mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^{*})$$

which is a linear fractional transformation, one has that

$$(4\text{-}8) \quad \leftrightarrow \quad \left|\left|\left[\left\{ \mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^{*}) - \frac{1}{\gamma}Q \right\}(I - \tilde{L}^{*}\tilde{L})^{-1/2} \ \ L\right]\right|\right|_{\infty} \leq 1 \tag{4-9}$$

Using the dual form of Theorem 1, the inequality 4-9 is equivalent to

$$\left|\left|(I - LL^{*})^{-1/2}\left\{ \mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^{*}) - \frac{1}{\gamma}Q \right\}(I - \tilde{L}^{*}\tilde{L})^{-1/2}\right|\right|_{\infty} \leq 1 .$$

QED

Remarks

(i)   This theorem is a generalization of Corollary 3.1.7 which was proved in [D1]. The difference is that the causality is a constraint in the problem considered h·.e.

(ii) S and $\bar{S}$ need not to be spectral factors. S and $\bar{S}$ can be any square transfer function matrices such that $S^{*}S = (\gamma^2 I - R_{22}^{*}R_{22})$ and $\bar{S}\bar{S}^{*} = (\gamma^2 I - R_{22}R_{22}^{*})$.

(iii)   The linear fractional transformation $F_l(\frac{1}{\gamma}R,\frac{1}{\gamma}R\frac{*}{22})$ has no poles on the $j\omega$-axis. This is true

because $R \in L_\infty$ and $\gamma > ||R_{22}||_\infty$.

The next theorem provides the upper and lower bounds for the γ-iteration of 4-block **GDP**.

Theorem 4

Let

$$\gamma_o = \min_{Q \in H_\infty} \left|\left|\begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\right|\right|_\infty$$

$$\gamma_{u_1} = \min \left\{ \left|\left|\begin{bmatrix} ||\mathbf{H}_{R_{11}}|| \\ ||R_{12}||_\infty \\ ||[R_{21}\ R_{22}]||_\infty \end{bmatrix}\right|\right|_2 , \left|\left|\begin{bmatrix} ||\mathbf{H}_{R_{11}}|| \\ ||R_{21}||_\infty \\ \left|\left|\begin{bmatrix} R_{21} \\ R_{22} \end{bmatrix}\right|\right|_\infty \end{bmatrix}\right|\right|_2 \right\}$$

$$\gamma_{u_2} = \left|\left|\begin{bmatrix} ||\mathbf{H}_{R_{11}}|| & ||R_{12}||_\infty \\ ||R_{21}||_\infty & ||R_{22}||_\infty \end{bmatrix}\right|\right|_F$$

$$\gamma_{l_1} = \max \left\{ ||\mathbf{H}_R|| , ||[R_{21}\ R_{22}]||_\infty , \left|\left|\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}\right|\right|_\infty \right\}$$

$$\gamma_{l_2} = \max \left\{ \left|\left|\mathbf{H}_{\begin{bmatrix} R_{11} \\ R_{21} \end{bmatrix}}\right|\right| , ||\mathbf{H}_{[R_{11}\ R_{12}]}|| , ||[R_{21}\ R_{22}]||_\infty , \left|\left|\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}\right|\right|_\infty \right\}$$

and

$$\gamma_{l_3} = \max \left\{ ||\mathbf{H}_{R_{11}}|| , ||R_{21}||_\infty , ||R_{12}||_\infty , ||R_{22}||_\infty \right\}$$

Then

$$\frac{1}{2}\gamma_o \leq \gamma_{l_3} \leq \gamma_{l_2} \leq \gamma_{l_1} \leq \gamma_o \leq \gamma_{u_1} \leq \gamma_{u_2} \leq 2\gamma_o .$$

Proof

Let Q_{opt}, $\hat{Q}_{opt}$, $\tilde{Q}_{opt}$ satisfy

$$\left\| \begin{bmatrix} R_{11} - Q_{opt} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty = \min_{Q \in H_\infty} \left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty$$

$$\| [R_{11} - \hat{Q}_{opt} \ R_{12}] \|_\infty = \min_{\hat{Q} \in H_\infty} \| [R_{11} - \hat{Q} \ R_{12}] \|_\infty$$

and

$$\| R_{11} - \tilde{Q}_{opt} \|_\infty = \min_{\tilde{Q} \in H_\infty} \| R_{11} - \tilde{Q} \|_\infty$$

The proof is broken into two parts:

(A) Upper bounds $(\gamma_o \leq \gamma_{u_1} \leq \gamma_{u_2} \leq 2 \gamma_o)$.

To establish that $\gamma_o \leq \gamma_{u_1}$, note that

$$\left\| \begin{bmatrix} R_{11} - Q_{opt} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty \leq \left\| \begin{bmatrix} R_{11} - \hat{Q}_{opt} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty$$

$$\leq \left\| \begin{bmatrix} \| [R_{11} - \hat{Q}_{opt} \ R_{12}] \|_\infty \\ \| [R_{21} \ R_{22}] \|_\infty \end{bmatrix} \right\|_2 \qquad \text{(from Lemma A.1)}$$

$$\leq \left\| \begin{bmatrix} \| R_{11} - \tilde{Q}_{opt} \|_\infty \\ \| R_{21} \|_\infty \\ \left\| \begin{bmatrix} R_{21} \\ R_{22} \end{bmatrix} \right\|_\infty \end{bmatrix} \right\|_2 \qquad \text{(from Theorem 2)}$$

$$\therefore \qquad \gamma_o \leq \gamma_{u_1} .$$

That $\gamma_{u_1} \leq \gamma_{u_2}$ is obvious.

To demonstrate that $\gamma_{u_2} \leq 2 \gamma_o$, note that

$$\left\| \begin{bmatrix} \| R_{11} - \tilde{Q}_{opt} \|_\infty & \| R_{12} \|_\infty \\ \| R_{21} \|_\infty & \| R_{22} \|_\infty \end{bmatrix} \right\|_F \leq \left\| \begin{bmatrix} \| R_{11} - Q_{opt} \|_\infty & \| R_{12} \|_\infty \\ \| R_{21} \|_\infty & \| R_{22} \|_\infty \end{bmatrix} \right\|_F$$

$$\leq 2 \left|\left|\begin{bmatrix} R_{11} - Q_{opt} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\right|\right|_{\infty} \qquad \text{(from Lemma A.2)}$$

$$\therefore \qquad \gamma_{u_2} \leq 2\gamma_o \ .$$

(B) Lower bounds $\ (\ \frac{1}{2} \gamma_o \ \leq \gamma_{l_3} \leq \gamma_{l_2} \leq \gamma_{l_1} \leq \gamma_o \ )$.

To show that $\gamma_{l_1} \leq \gamma_o$, note that

$$||\mathbf{H}_R|| = \min_{Q \in H_{\infty}} ||R - Q||_{\infty} \leq \left|\left| R - \begin{bmatrix} Q_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} = \gamma_o \ .$$

In addition, it is easy to see that

$$||[R_{21} \ R_{22}]||_{\infty} \leq \gamma_o \quad \text{and} \quad \left|\left|\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}\right|\right|_{\infty} \leq \gamma_o \ .$$

To show that $\gamma_{l_2} \leq \gamma_{l_1}$ note that both $[R_{21} \ R_{22}]$ and $\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}$ are the "compression" of R , hence, the inequality follows immediately.

Since R_{11} is a "compression" of $[R_{11} \ R_{12}]$ and $\begin{bmatrix} R_{11} \\ R_{21} \end{bmatrix}$, and

$$\max \left\{ ||R_{21}||_{\infty}, ||R_{12}||_{\infty}, ||R_{22}||_{\infty} \right\} \leq \max \left\{ ||[R_{21} \ R_{22}]||_{\infty}, \left|\left|\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix}\right|\right|_{\infty} \right\} \ ,$$

$$\therefore \qquad \gamma_{l_3} \leq \gamma_{l_2} \ .$$

Finally, to demonstrate that $\frac{1}{2} \gamma_o \ \leq \gamma_{l_3}$ note that

$$\gamma_o \ \leq \gamma_{u_2} \leq 2 \max \left\{ ||\mathbf{H}_{R_{11}}||, ||R_{21}||_{\infty}, ||R_{12}||_{\infty}, ||R_{22}||_{\infty} \right\} = 2 \gamma_{l_3}$$

$$\therefore \qquad \frac{1}{2} \gamma_o \ \leq \gamma_{l_3} \ .$$

QED

Remarks

(i)   Note that $||R||_\infty = \left|\left|\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\right|\right|_\infty$ is also an upper bound.

(ii)   The proof of Theorem 4 implies that the γ-iteration may be avoided altogether by simply using $\bar{Q}_{opt}$ instead of Q_{opt}, with the guarantee that the solution will be suboptimal by no more than a factor of 2. Note that $\bar{Q}_{opt}$ is simply the best approximation to R_{11}. For the 2-block case this approximation method will be suboptimal by no more than $\sqrt{2}$.

This section will be ended with a description for the γ-iteration procedure. Both 2-block and 4-block cases will be discussed.

I. 2-block case (γ-Iteration procedure) :

(i)   Compute the lower bound $||R_2||_\infty$.

(ii)   Choose γ such that $||R_2||_\infty < \gamma$.

(iii) Find a spectral factor $M = (\gamma^2 I - R_2^* R_2)^{1/2}$.

(iv)   Let $f(\gamma) = ||\mathbf{H}_{R_1 M^{-1}}||$.

       (a) if $f(\gamma) > 1$, go to (ii) and choose a larger γ.

       (b) if $f(\gamma) < 1$ and $\gamma > ||R_2||_\infty$, go to (ii) and choose a smaller γ.

       (c) if $f(\gamma) = 1$, go to (v).

(v)   The value of γ is the minimal achievable norm. Find a best approximation of $R_1 M^{-1}$, denoted by $\bar{Q}_{opt}$.

(vi)   An optimal solution is $Q_{opt} = \bar{Q}_{opt} M$.

II. 4-block case (γ-iteration procedure) :

(i)   Compute the lower bound $\gamma_l = \max \left\{ \; ||[R_{21} \; R_{22}]||_\infty \; , \; \left\| \begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix} \right\|_\infty \right\}$.

(ii)   Choose γ such that $\gamma > \gamma_l$.

(iii)   Find the spectral factor $S = (\gamma^2 I - R_{22}^* R_{22})^{1/2}$ and $\bar{S} = (\gamma^2 I - R_{22} R_{22}^*)^{1/2}$. Let $L = R_{12} S^{-1}$ and $\bar{L} = \bar{S}^{-1} R_{21}$.

(iv)   Find the spectral factor $(I - LL^*)^{1/2}$ and $(I - \bar{L}^* \bar{L})^{1/2}$.

(v)   Let $f(\gamma) = \left\| \mathbf{H}_{(I-LL^*)^{-1/2} \mathbf{F}_l (\frac{1}{\gamma} R , \frac{1}{\gamma} R_{22}^*)(I-\bar{L}^* \bar{L})^{-1/2}} \right\|$ where $\mathbf{F}_l(\frac{1}{\gamma} R , \frac{1}{\gamma} R_{22}^*)$ is a linear fractional transformation,

(a) if $f(\gamma) > 1$, go to (ii) and choose a larger γ.

(b) if $f(\gamma) < 1$, go to (ii) and choose a smaller γ.

(c) if $f(\gamma) = 1$, go to (vi).

(vi)   The value of γ is the minimal achievable norm. Find a best approximation of $(I - LL^*)^{-1/2} \mathbf{F}_l(\frac{1}{\gamma} R , \frac{1}{\gamma} R_{22}^*)(I - \bar{L}^* \bar{L})^{-1/2}$, denoted by $\tilde{Q}_{opt}$.

(vii)   An optimal solution is $Q_{opt} = \gamma(I - LL^*)^{1/2} \tilde{Q}_{opt} (I - \bar{L}^* \bar{L})^{1/2}$.

The algorithms are not complete without some method for selecting the next guess for γ in step (iv) (2-block) or step (v) (4-block). The guaranteed convergence rate for the algorithm will depend on this method and what can be proven about the relationship between γ and $f(\gamma)$. This relationship is the focus of the next section. The justification of the above algorithms can also be seen there.

4.2 Properties of γ-iteration

This section develops some properties of the γ-iteration in the 2-block **GDP** that are useful in selecting the next guess for γ in step (iv) of the algorithm in the previous section. The guaranteed convergence rate for the algorithm will depend on this method and what can be proven about the relationship between γ and $\bar\gamma$.

Recall from Theorem 1 in the last section that for a given γ, $\bar\gamma$ can be found as the norm of the Hankel operator $H_{R_1 M^{-1}}$. Since R_1 and M are rational, this operator is finite rank and $\bar\gamma$ can be found as the square root of the largest eigenvalue λ_{max} of the following standard eigenvalue problem:

$$(H_{R_1 M^{-1}})^*(H_{R_1 M^{-1}})u \;=\; \lambda u \tag{SEP}$$

or equivalently as the square root of the maximum eigenvalue of the following generalized eigenvalue problem:

$$H_{R_1}{}^* H_{R_1} v \;=\; \lambda(\gamma^2 I - T_{R_2^* R_2})\, v \tag{GEP}$$

To see this equivalence note that since $M^{-1} \in H_\infty$, $H_{R_1 M^{-1}}$ is equal to $H_{R_1} T_{M^{-1}}$. Hence,

$$(\text{SEP}) \;\Longleftrightarrow\; T_{M^{-1}}{}^* H_{R_1}{}^* H_{R_1} T_{M^{-1}} u \;=\; \lambda u$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda (T_{M^{-1}}{}^*)^{-1}(T_{M^{-1}})^{-1} v \qquad (v = T_{M^{-1}} u)$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda T_M{}^* T_M v$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda T_{M^*} T_M v$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda T_{M^* M}\, v \qquad (\text{Sarason (1967)})$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda (T_{\gamma^2 I - R_2^* R_2}) v$$

$$\Longleftrightarrow\; H_{R_1}{}^* H_{R_1} v \;=\; \lambda(\gamma^2 I - T_{R_2^* R_2}) v \tag{GEP}$$

This is the desired generalized eigenvalue problem (GEP) and the dependency of the (generalized) eigenvalues on γ is clear. Because λ and v in (GEP) can be obtained from the standard eigenvalue problem in (SEP), existence of eigenvalues or eigenvectors for (GEP) is assured. The eigenvalues of (GEP) are functions of γ and are nonnegative for all $\gamma > ||T_{R_2^* R_2}||^{1/2} = ||R_2||_\infty$. We shall prove that $\lambda_{max}(\gamma)$ (and its square root) is continuous, strictly monotonically decreasing and convex in γ where, for a given

$\gamma > ||R_2||_\infty$, $\lambda_{max}(\gamma)$ is defined as the maximum eigenvalue of (GEP). Results stated as Theorem 7 show that $\lambda_{max}(\gamma)$ can be bounded by simple functions. These properties can be used to guess successive γ's in a way that should quickly converge to the optimal norm, γ_o. Theorems 8 and 9 give the conditions for which the minimal norm is achieved.

A key observation of (GEP) is that it can be regarded as a "perturbed" generalized eigenvalue problem. Therefore, the perturbation theory of generalized eigenvalue problems for a special case which can later be applied to (GEP) will be considered first. The results can then be used to prove the properties mentioned above.

Consider the following generalized eigenvalue problem for Hilbert space operators A and B,

$$Av(t) = \lambda(t)B(t)v(t) \qquad t \in (-\epsilon,\epsilon) , \quad \epsilon > 0 \qquad \text{(GEP1)}$$

where

A is finite rank, positive semi-definite and independent of t,

and

$B(t)$ is bounded, positive-definite and analytic in the neighborhood of $t = 0$.

Since A and $B(t)$ are self-adjoint and $B(t) > 0$, it is well-known (Kato (1976)) that by appropriate ordering of the eigenvalues $\{\lambda_i\}$ and selection of eigenvectors $\{v_i\}$, it is possible to pair eigenvalues and eigenvectors $\{\lambda_i(t), v_i(t)\}$ such that

$$Av_i(t) = \lambda_i(t)B(t)v_i(t)$$

for all t, i and $\{\lambda_i(t)\}, \{v_i(t)\}$ are analytic for all $t \in (-\epsilon,\epsilon)$. At values of t where (GEP1) has simple eigenvalues, this is trivial. At degenerate points, it requires the selection of $\lambda_i(t)$, $v_i(t)$ such that the analyticity is retained through (isolated) point where eigenvalues coalesce. Therefore, $\lambda(t)$, $v(t)$ and $B(t)$ can be written in Taylor series expansion as follows:

$$\lambda(t) = \lambda_0 + t\dot{\lambda}_0 + \frac{t^2}{2}\ddot{\lambda}_0 + \ldots\ldots \quad (\lambda_0 = \lambda(0), \dot{\lambda}_0 = \dot{\lambda}(0), \ldots)$$

$$v(t) = v_0 + t\dot{v}_0 + \frac{t^2}{2}\ddot{v}_0 + \ldots\ldots \quad (v_0 = v(0), \dot{v}_0 = \dot{v}(0), \ldots)$$

(4-10a)

and

$$B(t) = B_0 + t\dot{B}_0 + \frac{t^2}{2}\ddot{B}_0 + \ldots\ldots \quad (B_0 = B(0), \dot{B}_0 = \dot{B}(0), \ldots)$$

Substitution of these expressions into (GEP1) yields

$$A\left(v_0 + t\dot{v}_0 + \frac{t^2}{2}\ddot{v}_0 + \ldots\right)$$

$$= \left(\lambda_0 + t\dot{\lambda}_0 + \frac{t^2}{2}\ddot{\lambda}_0 + \ldots\right)\left(B_0 + t\dot{B}_0 + \frac{t^2}{2}\ddot{B}_0 + \ldots\right)\left(v_0 + t\dot{v}_0 + \frac{t^2}{2}\ddot{v}_0 + \ldots\right)$$

(4-11)

$$\longleftrightarrow$$

$$Av_0 + tA\dot{v}_0 + \frac{t^2}{2}A\ddot{v}_0 + \ldots = (\lambda_0 B_0 v_0) + t(B_0 + \dot{\lambda}_0 B_0 v_0 + \lambda_0 \dot{B}_0 v_0 + \lambda_0 B_0 \dot{v}_0)$$

$$+ \frac{t^2}{2}(\ddot{\lambda}_0 B_0 v_0 + \lambda_0 \ddot{B}_0 v_0 + \lambda_0 B_0 \ddot{v}_0 + 2\dot{\lambda}_0 \dot{B}_0 v_0$$

$$+ 2\lambda_0 \dot{B}_0 \dot{v}_0 + 2\dot{\lambda}_0 B_0 \dot{v}_0) + \ldots\ldots$$

(4-12)

Eq. 4-12 is true $\forall\ t \in (-\epsilon, \epsilon)$. This implies

(i) t^0 (constant) term:

$$Av_0 = \lambda_0 B_0 v_0$$

(4-13a)

(ii) t^1 term:

$$A\dot{v}_0 = \dot{\lambda}_0 B_0 v_0 + \lambda_0 \dot{B}_0 v_0 + \lambda_0 B_0 \dot{v}_0$$

(4-13b)

(iii) t^2 term:

$$A\ddot{v}_0 = \ddot{\lambda}_0 B_0 v_0 + \lambda_0 \ddot{B}_0 v_0 + \lambda_0 B_0 \ddot{v}_0 + 2(\dot{\lambda}_0 \dot{B}_0 v_0 + \lambda_0 \dot{B}_0 \dot{v}_0 + \dot{\lambda}_0 B_0 \dot{v}_0)$$

(4-13c)

For simplicity, λ_0 has multiplicity 1 and $t = 0$ is assumed to be a regular point, i.e., multiplicity of $\lambda(t)$ is constant $\forall\ t \in (-\epsilon, \epsilon)$.

Note that $\{v_i(t)\}$ form a complete set of eigenvectors. Hence, the generalized eigenvalue problem considered here can also be expressed as : find $\Lambda(t)$ such that

$$AV(t) = B(t)V(t)\Lambda(t) \quad \text{(for some } V(t) \neq 0) \quad \text{(GEP2)}$$

where

$$V(t) = \begin{bmatrix} v(t) & V_\perp(t) \end{bmatrix} \quad \text{and} \quad \Lambda(t) = \begin{bmatrix} \lambda(t) & 0 \\ 0 & \Lambda_\perp(t) \end{bmatrix}.$$

Lemma 1

Assume $V_\perp = V_\perp(0)$ and v_o and B_o as defined in Eqs. 4-10b and 4-10c, then

$$v_o^* B_o V_\perp = 0.$$

Proof

Consider (GEP2) at $t = 0$,

$$AV_o = B_o V_o \Lambda_o$$

$$\Longrightarrow \quad [(B_o^{-1/2})^* A B_o^{-1/2}](B_o^{1/2} V_o) = (B_o^{1/2} V_o)\Lambda_o \quad (4-14)$$

where $B_o^{1/2}$ is any nonsingular matrix such that $(B_o^{1/2})^* B_o^{1/2} = B_o$.

Finding Λ_o to satisfy Eq. 4-14 is a standard eigenvalue problem where $(B_o^{-1/2})^* A (B_o)^{-1/2}$ is Hermitian (in fact, nonnegative). It is well-known that the eigenvectors of a Hermitian matrix, if they correspond to different eigenvalues, are orthogonal to one another. Therefore,

$$(B_o^{1/2} v_o)^* B_o^{1/2} V_\perp = v_o^* B_o V_\perp = 0.$$

<div align="right">QED</div>

The following theorem gives the formula for the first and second derivative of the eigenvalue of (GEP1) $\lambda(t)$ at $t = 0$.

Theorem 2

Consider the (GEP1) with Taylor series expansion as given in Eqs 4-10a, 4-10b, and 4-10c. Then

(i) $\quad \dot{\lambda}_o = -\lambda_o \dfrac{v_o^* \dot{B} v_o}{v_o^* B_o v_o}$

(ii) $\quad \ddot{\lambda}_\circ \;=\; -\lambda_\circ\,\dfrac{v_\circ{}^*\ddot{B}_\circ\,v_\circ}{v_\circ{}^*B_\circ\,v_\circ}\;+\;2\lambda_\circ\,(\dfrac{v_\circ{}^*\dot{B}_\circ v_\circ}{v_\circ{}^*B_\circ\,v_\circ})^2$

$$+\;\dfrac{2\lambda_\circ{}^2}{v_\circ{}^*B_\circ\,v_\circ}(V_\perp{}^*\;\dot{B}_\circ v_\circ)^*[(\lambda_\circ I-\Lambda_\perp)(V_\perp{}^*\;B_\circ\;V_\perp)]^{-1}(V_\perp{}^*\;\dot{B}_\circ v_\circ)$$

Proof

(i) To get the expression for $\dot{\lambda}_\circ$, consider

$v_\circ{}^*\bullet$ (Eq. 4-13b) $\quad ==>$

$$v_\circ{}^*A\dot{v}_\circ \;=\; \dot{\lambda}_\circ v_\circ{}^*B_\circ\,v_\circ \;+\; \lambda_\circ\,v_\circ{}^*\dot{B}_\circ v_\circ \;+\; \lambda_\circ\,v_\circ{}^*B_\circ\,\dot{v}_\circ$$

$$(Av_\circ-\lambda_\circ B_\circ v_\circ)^*\dot{v}_\circ \;=\; \dot{\lambda}_\circ v_\circ{}^*B_\circ\,v_\circ \;+\; \lambda_\circ\,v_\circ{}^*\dot{B}_\circ v_\circ$$

$$0 \;=\; \dot{\lambda}_\circ v_\circ{}^*B_\circ\,v_\circ \;+\; \lambda_\circ\,v_\circ{}^*\dot{B}_\circ v_\circ$$

$$\therefore \qquad \dot{\lambda}_\circ \;=\; -\lambda_\circ\,\dfrac{v_\circ{}^*\dot{B}_\circ v_\circ}{v_\circ{}^*B_\circ\,v_\circ}$$

(ii) To get the expression for $\ddot{\lambda}_\circ$ consider

$V_\perp{}^*\;\bullet$ (Eq. 4-13b) $\quad ==>$

$$V_\perp{}^*\,A\dot{v}_\circ \;=\; \dot{\lambda}_\circ V_\perp{}^*\,B_\circ\,v_\circ \;+\; \lambda_\circ\,V_\perp{}^*\,\dot{B}_\circ v_\circ \;+\; \lambda_\circ\,V_\perp{}^*\,B_\circ\,\dot{v}_\circ$$

From Lemma 1,

$$(AV_\perp-\lambda_\circ B_\circ V_\perp)^*\dot{v}_\circ \;=\; \lambda_\circ\,V_\perp{}^*\,\dot{B}_\circ v_\circ$$

$$(B_\circ V_\perp\Lambda_\perp-\lambda_\circ B_\circ V_\perp)^*\dot{v}_\circ \;=\; \lambda_\circ\,V_\perp{}^*\,\dot{B}_\circ v_\circ$$

$$(\Lambda_\perp-\lambda_\circ I)(B_\circ V_\perp)^*\dot{v}_\circ \;=\; \lambda_\circ\,V_\perp{}^*\,\dot{B}_\circ v_\circ \qquad\qquad (4\text{-}15)$$

Since $v_\circ$ and the columns of $V_\perp$ form a basis, $\dot{v}_\circ$ can be expressed as:

$$\dot{v}_\circ \;=\; \alpha v_\circ \;+\; V_\perp z \qquad\qquad (4\text{-}16)$$

where α is a scalar and z is a vector. Therefore,

$$(B_\circ V_\perp)^*\dot{v}_\circ \;=\; V_\perp{}^*\,B_\circ\,(\alpha v_\circ+V_\perp z)$$

$$=\; \alpha V_\perp{}^*B_\circ\,v_\circ \;+\; V_\perp{}^*\,B_\circ\,V_\perp z$$

$$=\; (V_\perp{}^*B_\circ V_\perp)z \qquad\qquad (4\text{-}17)$$

Substitution of the expression of Eq. 4-17 into Eq. 4-15 yields

$$(\Lambda_\perp - \lambda_o I)(V_\perp {}^* B_o V_\perp)z = \lambda_o V_\perp {}^* \dot{B}_o v_o$$

$$\therefore \quad z = \lambda_o [(\Lambda_\perp - \lambda_o I)(V_\perp {}^* B_o V_\perp)]^{-1} V_\perp {}^* \dot{B}_o v_o . \qquad (4\text{-}18)$$

Also consider

$v_o {}^* \bullet$ (Eq. 4-13c) $\implies$

$$v_o {}^* A \ddot{v}_o = \ddot{\lambda}_o v_o {}^* B_o v_o + \lambda_o v_o {}^* \ddot{B}_o v_o + \lambda_o v_o {}^* B_o \ddot{v}_o$$
$$+ 2(\dot{\lambda} v_o {}^* \dot{B}_o v_o + \lambda_o v_o {}^* \dot{B}_o \dot{v}_o + \dot{\lambda} v_o {}^* B_o \dot{v}_o)$$

$$(A v_o - \lambda_o B_o v_o){}^* \ddot{v}_o = \ddot{\lambda}_o v_o {}^* B_o v_o + \lambda_o v_o {}^* \ddot{B}_o v_o$$
$$+ 2(\dot{\lambda} v_o {}^* \dot{B}_o v_o + \lambda_o v_o {}^* \dot{B}_o \dot{v}_o + \dot{\lambda} v_o {}^* B_o \dot{v}_o)$$

$$-\ddot{\lambda}_o v_o {}^* B_o v_o = \lambda_o v_o {}^* \ddot{B}_o v_o + 2(\dot{\lambda} v_o {}^* \dot{B}_o v_o + \lambda_o v_o {}^* \dot{B}_o \dot{v}_o + \dot{\lambda} v_o {}^* B_o \dot{v}_o) \qquad (4\text{-}19)$$

Finally, substitution of the expressions of $\dot{\lambda}$, $\dot{v}_o$ and z into Eq. 4-19 yields

$$\ddot{\lambda}_o (v_o {}^* B_o v_o) = -\lambda_o v_o {}^* \ddot{B}_o v_o + 2\lambda_o \frac{(v_o {}^* \dot{B}_o v_o)^2}{v_o {}^* B_o v_o}$$
$$+ 2\lambda_o^2 (V_\perp {}^* \dot{B}_o v_o){}^* [(\lambda_o I - \Lambda_\perp)(V_\perp {}^* B_o V_\perp)]^{-1}(V_\perp {}^* \dot{B}_o v_o) .$$

QED

Remark

Note that the assumption that $t = 0$ is a regular point does not cause any problem. This is so because $\lambda(t)$ is analytic and $\dot{\lambda}(t)$ ($\ddot{\lambda}(t)$) must be continuous and $\dot{\lambda}_o (\ddot{\lambda}_o)$ can still be computed by taking the limit of $\dot{\lambda}(t)$ ($\ddot{\lambda}(t)$) as $t \to 0$.

Define $\sigma(t) = \lambda^{1/2}(t)$, $\sigma_o = \sigma(0)$, $\dot{\sigma}_o = \dot{\sigma}(0)$ and $\ddot{\sigma}_o = \ddot{\sigma}(0)$, then one has the expression for $\dot{\sigma}_o$ and $\ddot{\sigma}_o$, which are given in the following corollary.

Corollary 3

(i) $\quad \dot{\sigma}_\circ = -\dfrac{\sigma_\circ}{2}\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}$

(ii) $\quad \ddot{\sigma}_\circ = \dfrac{1}{2}\Big\{-\sigma_\circ\dfrac{v_\circ^*\ddot{B}_\circ\, v_\circ}{v_\circ^*B_\circ v_\circ} + \dfrac{3}{2}\sigma_\circ\big(\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}\big)^2$

$$+ \dfrac{2\sigma_\circ^3}{v_\circ^*B_\circ v_\circ}(V_\perp^*\ \dot{B}_\circ v_\circ)^*[(\lambda_\circ I - \Lambda_\perp)(V_\perp^*\ B_\circ V_\perp)]^{-1}(V_\perp^*\ \dot{B}_\circ v_\circ)\Big\}$$

Proof

Let

$$\sigma(t) = \sigma_\circ + t\dot{\sigma}_\circ + \dfrac{t^2}{2}\ddot{\sigma}_\circ + \dots\dots \quad (\sigma_\circ = \sigma(0),\ \dot{\sigma}_\circ = \dot{\sigma}(0), \dots)$$

$$\implies \quad \sigma^2(t) = \sigma_\circ^2 + t(2\sigma_\circ\dot{\sigma}_\circ) + \dfrac{t^2}{2}(2\sigma_\circ\ddot{\sigma}_\circ + 2\dot{\sigma}_\circ^2) + \dots\dots$$

$\therefore$

(a) $\quad \sigma_\circ^2 = \lambda_\circ$

(b) $\quad 2\sigma_\circ\dot{\sigma}_\circ = \dot{\lambda}_\circ$

and

(c) $\quad 2\sigma_\circ\ddot{\sigma}_\circ + 2\dot{\sigma}_\circ^2 = \ddot{\lambda}_\circ$

From (b) and Theorem 2-(i),

$$\dot{\sigma}_\circ = -\dfrac{\lambda_\circ}{2\sigma_\circ}\big(\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}\big) = -\dfrac{\sigma_\circ}{2}\big(\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}\big)$$

From (c) and Theorem 2-(ii),

$$\ddot{\sigma}_\circ = \dfrac{1}{2\sigma_\circ}(\ddot{\lambda}_\circ - 2\dot{\sigma}_\circ^2)$$

$$= \dfrac{1}{2\sigma_\circ}\Big\{-\sigma_\circ^2\dfrac{v_\circ^*\ddot{B}_\circ\, v_\circ}{v_\circ^*B_\circ v_\circ} + 2\sigma_\circ^2\big(\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}\big)^2 + \dfrac{2\sigma_\circ^4}{v_\circ^*B_\circ v_\circ}(V_\perp^*\ \dot{B}_\circ v_\circ)^*$$

$$[(\lambda_\circ I - \Lambda_\perp)(V_\perp^*\ B_\circ V_\perp)]^{-1}(V_\perp^*\ \dot{B}_\circ v_\circ) - \dfrac{\sigma_\circ^2}{2}\big(\dfrac{v_\circ^*\dot{B}_\circ v_\circ}{v_\circ^*B_\circ v_\circ}\big)^2\Big\}$$

$$= \frac{1}{2} \left\{ -\sigma_o \frac{v_o^* \dot{B}_o v_o}{v_o^* B_o v_o} + \frac{3}{2} \sigma_o (\frac{v_o^* \dot{B}_o v_o}{v_o^* B_o v_o})^2 + \frac{2\sigma_o^3}{v_o^* B_o v_o} (V_\perp^* \ \dot{B}_o v_o)^* \right.$$

$$\left. [(\lambda_o I - \Lambda_\perp)(V_\perp^* \ B_o V_\perp)]^{-1} (V_\perp^* \ \dot{B}_o v_o) \right\} .$$

QED

For $\gamma > ||R_2||_\infty$, define $\lambda_{max}(\gamma) \triangleq$ the maximum eigenvalue of (GEP) at a given γ and $\sigma_{max}(\gamma) \triangleq [\lambda_{max}(\gamma)]^{1/2}$.

The following two theorems are the main results of this section. First, λ_{max} and σ_{max} are shown to be strictly monotonically decreasing in γ.

Theorem 4

Consider the (GEP1) as given above. Then

(i) λ_{max} is continuous in γ.

(ii) σ_{max} is continuous in γ.

Proof

Let $t = \gamma - \bar{\gamma}$ and $B(t) = t^2 I + 2t\bar{\gamma}I + (\bar{\gamma}^2 I - T_{R_2^* R_2})$. The theorem follows immediately from the analyticity of eigenvalues with respect to t.

Theorem 5

Consider the (GEP1) as given above. Then

(i) λ_{max} is strictly monotonically decreasing in γ.

(ii) σ_{max} is strictly monotonically decreasing in γ.

Proof

Assume $\bar{\gamma} > ||R_2||_\infty$ is given.

To prove (i), let $t = \gamma - \bar{\gamma}$ and $B(t) = t^2 I + 2t\bar{\gamma}I + (\bar{\gamma}^2 I - T_{R_2^* R_2})$,

$$\therefore \quad \dot{B}_\bullet = 2\bar{\gamma}I$$

From Theorem 2-(i),

$$\dot{\lambda}_\bullet = -2\bar{\gamma}\lambda_\bullet \frac{v_\bullet^* v_\bullet}{v_\bullet^* B_\bullet v_\bullet} < 0$$

Therefore, λ_{max} is strictly monotonically decreasing with respect to t and hence, γ.

To prove (ii), define t the same way as in the proof of (i). From Corollary 3-(i),

$$\dot{\sigma}_\bullet = -\bar{\gamma}\sigma_\bullet \frac{v_\bullet^* v_\bullet}{v_\bullet^* B_\bullet v_\bullet} < 0$$

Therefore, σ_{max} is strictly monotonically decreasing with respect to γ.

Convexity is established next.

Theorem 6

Consider the (GEP1) as given above. Then

(i) λ_{max} is convex in γ.

(ii) σ_{max} is convex in γ.

Proof

The proofs will be given in a similar way to those in Theorems 4 and 5.

To establish (i) note that

$$\dot{B}_o = 2\bar{\gamma}I \quad \text{and} \quad \ddot{B}_o = 2I .$$

Since

$$-v_o{}^* \ddot{B}_o \ v_o \ + \ 2\frac{(v_o{}^* \dot{B}_\nu{}_o)^2}{v_o{}^* B_o \ v_o} \ = \ -2v_o{}^* v_o \ + \ 2(2\bar{\gamma})^2 \frac{(v_o{}^* v_o)^2}{v_o{}^* B_o \ v_o} \ = \ \frac{v_o{}^* v_o}{v_o{}^* B_o \ v_o}(8\bar{\gamma}^2 v_o{}^* v_o \ - \ 2v_o{}^* B_o \ v_o)$$

$$B_o \ = \ \bar{\gamma}^2 I \ - \ T_{R_1 R_2} \ < \ \bar{\gamma}^2 I$$

$$\Longrightarrow \quad \bar{\gamma}^2 v_o{}^* v_o \ > \ v_o{}^* B_o \ v_o$$

$$\Longrightarrow \quad 8\bar{\gamma}^2 v_o{}^* v_o \ - \ 2v_o{}^* B_o \ v_o \ > \ 0$$

If λ_o is the maximum eigenvalue , then $\ddot{\lambda}_o \ > \ 0$. It follows that λ_{max} is convex in γ.

To prove (ii) note that From Corollary 3-(ii),

$$\ddot{\sigma}_o \ = \ \frac{1}{2}\Big\{ -\sigma_o \frac{v_o{}^* \ddot{B}_o \ v_o}{v_o{}^* B_o \ v_o} \ + \ \frac{3}{2}\sigma_o (\frac{v_o{}^* \dot{B}_\nu{}_o}{v_o{}^* B_o \ v_o})^2 \ + \ \frac{2\sigma_o^3}{v_o{}^* B_o \ v_o}(V_\perp{}^* \ \dot{B}_\nu{}_o)^*$$

$$[(\lambda_o I - \Lambda_\perp)(V_\perp{}^* \ B_o \ V_\perp)]^{-1}(V_\perp{}^* \ \dot{B}_\nu{}_o)\Big\}$$

Since

$$-\frac{v_o{}^* \ddot{B}_o \ v_o}{v_o{}^* B_o \ v_o} \ + \ \frac{3}{2}\sigma_o (\frac{v_o{}^* \dot{B}_\nu{}_o}{v_o{}^* B_o \ v_o})^2 \ = \ \frac{v_o{}^* v_o}{v_o{}^* B_o \ v_o}(6\bar{\gamma}^2 v_o{}^* v_o \ - \ 2v_o{}^* B_o \ v_o) \ > \ 0$$

If λ_o is the maximum eigenvalue at $t = 0, \ddot{\sigma}_o \ > \ 0$. Therefore, σ_{max} is convex in γ.

QED

Remark

The idea of using generalized eigenvalue formulation (GEP) in this section is similar to that in Helton's broadband matching problem [H2] although the motivation here is completely different.

Although the function σ_{max} is unknown, the properties shown in Theorems 4 to 6 have provided some useful information about σ_{max} that can be used to obtain fast convergence of the γ-iteration. A detailed study of convergence rates is not given here, however some guidelines will be presented which

show how the properties of σ_{max} can be used to find a next guess for γ. One additional property of σ_{max} is useful in this regard and will be presented in the next theorem.

Define $\sigma_u(\bar{\gamma},\gamma) = \dfrac{c}{(\gamma^2 - \beta^2)^{1/2}}$ where $c = \sigma_{max}(\bar{\gamma})(\bar{\gamma}^2 - \beta^2)^{1/2}$ for some $\bar{\gamma} > \beta(= ||R_2||_\infty)$.

Theorem 7

Let σ_{max} and σ_u be defined as above. Then

(i) $\sigma_{max}(\gamma) < \sigma_u(\bar{\gamma},\gamma)$ if $\gamma < \bar{\gamma}$.

(ii) $\sigma_{max}(\gamma) = \sigma_u(\bar{\gamma},\gamma)$ if $\gamma = \bar{\gamma}$.

(iii) $\sigma_{max}(\gamma) > \sigma_u(\bar{\gamma},\gamma)$ if $\gamma > \bar{\gamma}$.

Proof

$$\sigma_u = \frac{c}{(\gamma^2 - \beta^2)^{1/2}}$$

$$\Longrightarrow \qquad \dot{\sigma}_u = -\frac{c\,\gamma}{(\gamma^2 - \beta^2)^{3/2}} = -\sigma_u \frac{\gamma}{\gamma^2 - \beta^2}$$

$$\therefore \qquad \frac{d}{d\gamma}(\ln \sigma_u) = -\frac{\gamma}{\gamma^2 - \beta^2}$$

From Corollary 3,

$$\dot{\sigma}_{max} = -\sigma_{max} \frac{\gamma v_o^* v_o}{v_o^*(\gamma^2 I - T_{R_2^* R_2}) v_o}$$

where v_o is the (generalized) eigenvector corresponding to $\lambda_{max}(\bar{\gamma})$.

$$\therefore \qquad \frac{d}{d\gamma}(\ln \sigma_{max}) = -\frac{\gamma v_o^* v_o}{v_o^*(\gamma^2 I - T_{R_2^* R_2}) v_o}$$

Since

$$(\gamma^2 - \beta^2)I \leq \gamma^2 I - T_{R_2^* R_2}$$

$$\implies \quad v_o{}^*(\gamma^2 - \beta^2)I v_o \;\leq\; v_o{}^*(\gamma^2 I - T_{R_2{}^*R_2}) v_o$$

$$\implies \quad \frac{1}{(\gamma^2 - \beta^2)} \;\geq\; \frac{v_o{}^* v_o}{v_o{}^*(\gamma^2 I - T_{R_2{}^*R_2}) v_o}$$

$$\therefore \quad \frac{d}{d\gamma}(\ln \sigma_u) \;\leq\; \frac{d}{d\gamma}(\ln \sigma_{max}) \;<\; 0 \tag{4-20}$$

Since $\sigma_{max}(\bar{\gamma}) = \sigma_u(\bar{\gamma},\bar{\gamma})$, Eq.(4-20) implies that

$$\ln(\sigma_u) \;\leq\; \ln(\sigma_{max}) \quad \text{if } \gamma > \bar{\gamma}$$

and

$$\ln(\sigma_u) \;\leq\; \ln(\sigma_{max}) \quad \text{if } \gamma < \bar{\gamma}$$

The theorem follows immediately.

QED

The importance of Theorem 7 can be seen from Figure 4-1. Suppose that at one step in the γ-iteration, we have evaluated σ_{max} at γ_l and γ_u from previous iterations, and want to make a new guess for γ. Assume that $\beta < \gamma_l < \gamma_u$ such that $\sigma_{max}(\gamma_l) > 1$ and $\sigma_{max}(\gamma_u) < 1$. From Theorem 7, we know immediately that $\gamma_l < \gamma_o < \gamma_u$. Since σ_{max} is a convex function in γ, σ_{max} must lie below the line segment (denoted by $F_1(\gamma)$) connecting the points $(\gamma_l, \sigma_{max}(\gamma_l))$ and $(\gamma_u, \sigma_{max}(\gamma_u))$. In addition, by Theorem 7, σ_{max} will lie above the function $\sigma_u(\gamma_l, \gamma) = \dfrac{\sigma_{max}(\gamma_l)(\gamma_l^2 - \beta^2)^{1/2}}{(\gamma^2 - \beta^2)^{1/2}}$ when $\gamma > \gamma_l$.

Suppose that $\gamma_u{}'$ and $\gamma_l{}'$ are the points where $F_1(\gamma_u{}') = 1$ and $\sigma_u(\gamma_l, \gamma_l{}') = 1$. We can conclude immediately that $\gamma_l{}' \leq \gamma_o \leq \gamma_u{}'$. The next guess for γ is narrowed considerably over what would be known on the basis of continuity, convexity, and monotonicity alone. Thus it is clearly possible to obtain a scheme for picking the next guess for γ that will provide rapid convergence to the optimal. Further consideration of convergence rates is beyond the scope of this report.

Remark

The γ-iteration can be viewed as the problem of finding the zero crossing of the function $(\sigma_{max}(\gamma) - 1)$.
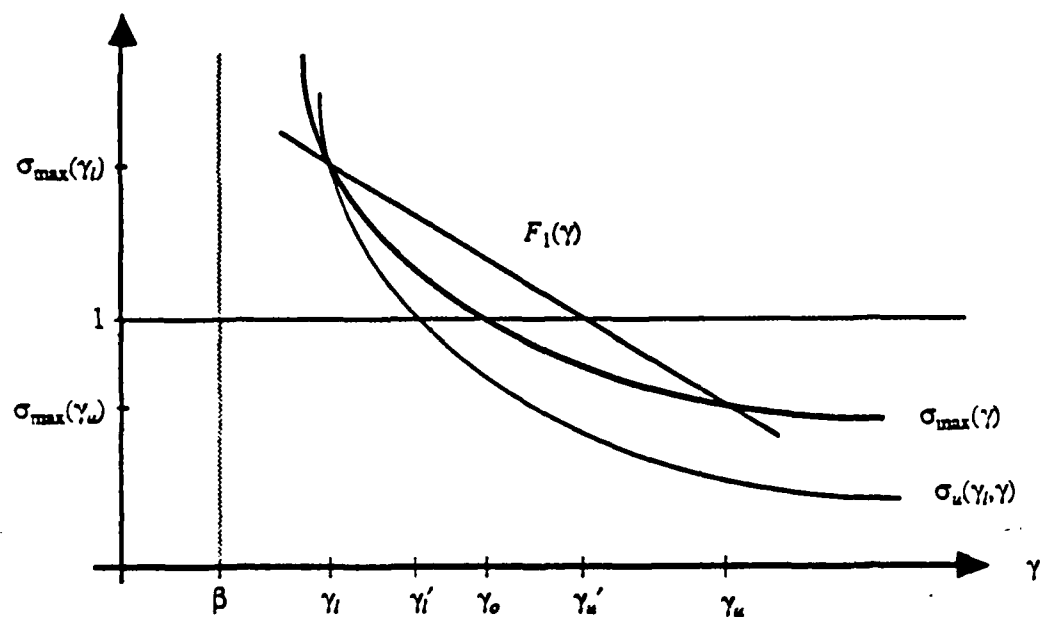

Figure 4-1. γ-iteration

The following two theorems give the conditions for which the optimal norm γ_o, is achieved.

Define $F(\gamma) = R_1(\gamma^2 I - R_2^* R_2)^{-1/2}$.

Theorem 8

Assume $\bar{\gamma} > ||R_2||_\infty$. Then

$$\min_{\tilde{Q} \in RH_\infty} ||F(\bar{\gamma}) - \tilde{Q}||_\infty = 1 \quad \text{if and only if} \quad \bar{\gamma} = \gamma_o .$$

Proof

(Necessity):

If $||H_{F(\bar{\gamma})}|| = 1$ but $\bar{\gamma} \neq \gamma_o$, then $\bar{\gamma}$ must be greater than γ_o (since γ_o is the minimum norm). From Theorem 5, $||H_{F(\bar{\gamma})}||$ is strictly monotonically decreasing with respect to γ. Therefore,

$$\bar{\gamma} > \gamma_o \quad ==> \quad ||H_{F(\gamma_o)}|| > ||H_{F(\bar{\gamma})}|| = 1$$

This contradicts the definition of γ_o. Hence, $\bar{\gamma} = \gamma_o$.

(Sufficiency):

If $\bar{\gamma} = \gamma_o$ but $||H_{F(\bar{\gamma})}|| < 1$, by continuity and monotonicity, $\exists \ \epsilon_o > 0$ such that

$$||H_{F(\tilde{\gamma})}|| \leq 1 \quad (\text{where } \tilde{\gamma} = \bar{\gamma} - \epsilon_o)$$

This is impossible since γ_o is the minimal achievable norm.

$$\therefore \quad ||H_{F(\bar{\gamma})}|| = 1.$$

<div align="right">QED</div>

Theorem 9

Assume $\bar{\gamma} = ||R_2||_\infty$. Then $\displaystyle\min_{\tilde{Q} \in RH_\infty} ||F(\bar{\gamma}) - \tilde{Q}|| \leq 1$ if and only if $\bar{\gamma} = \gamma_o$.

Proof

(Necessity) :

$||H_{F(\bar{\gamma})}|| \leq 1$ implies $\bar{\gamma} \geq \gamma_o$. Together with the lower bound (Theorem 4.1.2) $\gamma_o \geq ||R_2||_\infty$, therefore, $\gamma_o = \bar{\gamma}$.

(Sufficiency) :

This is obvious from the Theorem 4.1.1.

$\overline{\gamma} = ||R_2||_\infty$ and $||\mathbf{H}_{F(\overline{\gamma})}|| \leq 1$ if and only if $\overline{\gamma} = \gamma_o$.

QED

4.3 An Example

In this section, a simple example with a single parameter is constructed to illustrate various properties in the γ-iteration. An exact optimal solution will be derived in detail.

Consider the following 2-block problem:

$$\gamma_o = \min_{Q \in RH_\infty} \left\| \left[\begin{array}{c} \dfrac{1}{s-1} - Q \\ \dfrac{1}{s-a} \end{array} \right] \right\|_\infty \qquad (a > 0) \qquad (4\text{-}21)$$

Let Q_o be the optimal solution which achieves the minimum norm. Using the formula in Section 3,

$$\left\| \left[\begin{array}{c} \dfrac{1}{s-1} - Q \\ \dfrac{1}{s-a} \end{array} \right] \right\|_\infty \leq \gamma \qquad \Longleftrightarrow \qquad \left\| \left(\dfrac{1}{s-1} - Q \right) M^{-1} \right\|_\infty \leq 1$$

where $\quad M^{-1} = \left(\gamma^2 - \dfrac{1}{-s-a} \dfrac{1}{s-a} \right)^{-1/2} = \dfrac{s+a}{\gamma s + \sqrt{\gamma^2 a^2 - 1}}$.

Assume $\quad \bar{G} = \left(\dfrac{1}{s-1} M^{-1} \right)_{unstable} = \dfrac{\dfrac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}}}{s-1} \quad$ and consider the following best approximation problem:

$$\min_{\bar{Q} \in RH_\infty} \| \bar{G} - \bar{Q} \|_\infty .$$

It is not difficult to solve this problem. The minimum norm is

$$\min_{\bar{Q} \in RH_\infty} \| \bar{G} - \bar{Q} \|_\infty = \frac{1}{2} \left(\dfrac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}} \right) \qquad (4\text{-}22)$$

and the optimal solution is

$$\bar{Q}_o = -\frac{1}{2} \left(\dfrac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}} \right) . \qquad (4\text{-}23)$$

• Computation of the minimum norm γ_o :

Of course, in order to have Eq. 4-22 make sense, the Hankel norm of $\bar{G}$ must be less than or equal to 1. Therefore, it is reasonable to assume $\frac{1}{2}(\frac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}}) = 1$ and solve for γ. Once γ is found, substitution of it into Eqs. 4-22 and 4.23 leads to the optimal solution and minimum achievable norm of Eq. 4-21.

Let

$$\frac{1}{2}(\frac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}}) = 1$$

$$==>$$

$$1 + a = 2(\gamma + \sqrt{\gamma^2 a^2 - 1})$$

$$(1 + a) - 2\gamma = 2\sqrt{\gamma^2 a^2 - 1}$$

Taking square on both sides and collecting the terms with same power in γ yields

$$4(1 - a^2)\gamma^2 - 4(1 + a)\gamma + [4 + (1 + a)^2] = 0 \tag{4-24}$$

There are three cases to consider.

Case 1: $a = 1$

In this case, Eq. 4-24 is simply a linear equation and has solution $\gamma = 1$. Therefore,

$$\gamma_o = 1$$

If $a \neq 1$, Eq. (4) is a quadratic equation and the conditions such that a positive real solution exists are:

(i) $\qquad -\frac{4(1 + a)}{8(1 - a^2)} > 0 \qquad \Longleftrightarrow \qquad a > 1$

and

$$\left(\frac{1}{s-1}M^{-1}\right)_{stable} = \frac{-\dfrac{\gamma a - \sqrt{\gamma^2 a^2 - 1}}{\gamma + \sqrt{\gamma^2 a^2 - 1}}}{\gamma s + \sqrt{\gamma^2 a^2 - 1}}$$

Therefore,

$$Q_o = \left\{ (\tilde{Q}_o) + \left(\frac{1}{s-1}M^{-1}\right)_{stable} \right\} \cdot M^{-1} \Bigg|_{\gamma = \gamma_o}$$

$$= - \left\{ \frac{1}{2}\left(\frac{1+a}{\gamma + \sqrt{\gamma^2 a^2 - 1}}\right)\frac{\gamma s + \sqrt{\gamma^2 a^2 - 1}}{s+a} + \frac{\dfrac{\gamma a - \sqrt{\gamma^2 a^2 - 1}}{\gamma + \sqrt{\gamma^2 a^2 - 1}}}{s+a} \right\} \Bigg|_{\gamma = \gamma_o} \tag{4-25}$$

Substituting the optimal γ_o into Eq. 4-25 separately,

(i) $\quad a < 1 \quad : \quad Q_o = -\dfrac{(1+a)s + 2a}{2(s+a)}.$ $\hfill (4-26)$

(ii) $\quad a = 1 \quad : \quad Q_o = -1.$ $\hfill (4-27)$

(iii) $\quad a > 1 \quad : \quad Q_o = -\dfrac{\gamma_o s + \gamma_o + \dfrac{a-1}{2}}{s+a}$ $\hfill (4-28)$

where $\gamma_o = \dfrac{1}{2(a-1)}\left[-1 + \sqrt{a^2 + \dfrac{4(a-1)}{a+1}}\right]$

Remarks

The significance of this example can be stated as follows:

(i) From Eq. 4-22, the Hankel norm of $\tilde{G}$ is a convex function of γ. This can be verified by computing its second derivative with respect to γ and showing that it is always greater than zero.

(ii) If $0 < a \leq 1$, then $\gamma_o = ||R_2||_\infty$. This means that the lower bound in Theorem 4.1.2 is tight.

(iii) If $0 < a < 1$, then the optimal solution is not unique. Recently, Professor Bruce Francis showed a family of optimal solutions which includes the one given in Eq. 4-26 [F6]. However, if $a \geq 1$, it can be shown that the optimal solution (Eqs. 4-27 and 4-28) is unique.

(ii) $\quad [4(1+a)^2]^2 - 4[4(1-a^2)] \, [4 + (1+a)^2] \geq 0 \quad \Longleftrightarrow \quad a^3 + a^2 + 4a - 4 \geq 0$

Case 2: $a > 1$

It is easy to see that (i) and (ii) are satisfied automatically in this case. Hence, in (4) will have a real positive solution if $a > 1$ and the solution is

$$\gamma_s \; = \; \frac{1}{2(a-1)} \left[-1 + \sqrt{a^2 + \frac{4(a-1)}{a+1}} \, \right]$$

Case 3: $a < 1$

In this case, Eq. (4) has no positive real solution. But observe that if $0 < a < 1$

$$\frac{1}{2} \, (\frac{1+a}{\sqrt{\gamma + \gamma^2 a^2 - 1}}) \; < \; 1 \qquad \forall \; \gamma \geq \frac{1}{a}$$

where $\frac{1}{a} = || \, \frac{1}{s-a} \, ||_\infty$. Therefore, the smallest possible value of γ is simply $\frac{1}{a}$.

$$\therefore \qquad \gamma_s \; = \; \frac{1}{a}$$

To summarize,

(i) $\qquad a < 1 \quad : \quad \gamma_s \; = \; \frac{1}{a}$

(ii) $\qquad a = 1 \quad : \quad \gamma_s \; = \; 1$

(iii) $\qquad a > 1 \quad : \quad \gamma_s \; = \; \frac{1}{2(a-1)} \left[-1 + \sqrt{a^2 + \frac{4(a-1)}{a+1}} \, \right]$

• Recovery of the optimal solution:

Note that

CHAPTER 5

OPTIMAL STATE-SPACE SOLUTIONS OF
GENERAL DISTANCE PROBLEMS

In this chapter, the general distance problem is treated in the context of the real-rational transfer function matrix $(R(s) \in RL_\infty)$ where H_∞ of the half-plane instead of the unit disc is used. Both two-block and four-block problems are considered separately. Since any real-rational transfer function matrix has a finite-dimensional state-space representation, it is shown that the corresponding **GDP** can be solved using attractive state-space methods involving only standard real matrix operations.

Recall that the **GDP** can be simplified to an equivalent best approximation problem if the value of γ is chosen properly (Theorems 3.4.1 and 3.4.4) where the spectral factor of the para-Hermitian matrix with the form $(\gamma^2 I - G^*G)$ (or $(\gamma^2 I - GG^*)$) is used. Since the stability of G does not affect the result of the spectral factorization, i.e., the poles of G in the open rhp will be replaced by their counterpart in the open lhp, it would be convenient to assume $G \in RH_\infty$. If this is not the case, a coprime factorization of G with inner denominator can always be found and the numerator matrix can be used in place of G. This particular coprime factorization is presented in Section 5.1 which involves finding the stabilizing solution of an **ARE**. Since the constant term in this **ARE** is identically zero, the stabilizing solution can be obtained more efficiently by a "modified" Schur method which is no more than finding the Schur decomposition and solving a set of linear equations. This is discussed in Section 5.2. In Section 5.3, the state-space formula of spectral factors $(\gamma^2 I - G^*G)^{1/2}$ and $(\gamma^2 I - GG^*)^{1/2}$ is derived. Once the **GDP** is simplified to the corresponding best approximation problem, Glover's algorithm [G2] can be applied to find the solution which is reviewed in Section 5.4. Finally, combining the state-space formula of factorizations and the best approximation mentioned in Sections 5.1 through 5.4, the optimal solution for the **GDP** is obtained in terms of the "closed-form" state-space realization.

5.1 Coprime Factorizations with Inner Denominator

In this section, the coprime factorization with inner denominator will be developed. Explicit state-space realizations will be given which involves solving an algebraic Riccati equation. Without loss of generality, it is assumed that $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in R_p^{p \times m}$ and the realization is minimal.

Theorem 1

Consider $G \in R_p^{p \times m}$. Then there exists a rcf $G = NM^{-1}$ such that $M \in RH_\infty^{m \times m}$ is inner if and only if G has no poles on the $j\omega$-axis. A particular realization is

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array}\right] \in RH_\infty^{(m+p) \times m} \tag{5-1}$$

where

$$F = -B^T X \tag{5-2}$$

and

$$X = \mathbf{Ric} \begin{bmatrix} A & -BB^T \\ 0 & -A^T \end{bmatrix} \geq 0 . \tag{5-3}$$

Proof

(Sufficiency):

Assume G has no poles on the $j\omega$-axis, therefore, $\mathrm{Re}[\lambda_i(A)] \neq 0$. It is known that the rcf $G = NM^{-1}$ can be obtained using state-feedback gain F such that $(A+BF)$ is a stability matrix where M and N have the realization as in Eq. 5-1. Since M is required to be inner, the following two equations must be satisfied for some X (Lemma 2.3.3):

$$B^T X + F = 0 \tag{5-4}$$

and

$$(A+BF)^T X + X(A+BF) + F^T F = 0 . \tag{5-5}$$

Note that

(Eq. 5-4) $\Longleftrightarrow$ $F = -B^T X$

(Eq. 5-5) $\Longleftrightarrow$ $A^T X - XBB^T X + XA - XBB^T X + XBB^T X = 0$

$\Longleftrightarrow$ $A^T X + XA - XBB^T X = 0 . \tag{5-6}$

Eq. 5-6 is an algebraic Riccati equation. Since (A,B) is a controllable pair and $\text{Re}[\lambda, (A)] \neq 0$, Eq. 5-6 has a unique stabilizing solution $X \geq 0$.

(Necessity):

$$GG^* = NM^{-1}(NM^{-1})^* = NM^{-1}(M^{-1})^* N^* = NN^* .$$

Since $N \in RH_\infty$, GG^* has no poles on the $j\omega$-axis. Therefore, G has no poles on the $j\omega$-axis.

QED

Similar results for the lcf can be established easily by duality.

Corollary 2

There exists a lcf $G = \bar{M}^{-1}\bar{N}$ such that $\bar{M} \in RH_\infty^{p \times p}$ is inner if and only if G has no poles on the $j\omega$-axis. A particular realization is

$$\begin{bmatrix} \bar{M} & \bar{N} \end{bmatrix} = \left[\begin{array}{c|cc} A+HC & H & B+HD \\ \hline C & I & D \end{array} \right] \in RH_\infty^{p \times (p+m)} \tag{5-7}$$

where

$$H = -YC^T \tag{5-8}$$

and

$$Y = \text{Ric} \begin{bmatrix} A^T & -C^T C \\ 0 & -A \end{bmatrix} \geq 0 . \tag{5-9}$$

Remarks

(i)   The minimality condition on G is not necessary and can be weaken to that the pair (A, B) $((C, A))$ is stabilizable (detectable) and $\text{Re}[\lambda_i(A)] \neq 0$ in Theorem 1 (Corollary 2).

(ii)   Note that the constant term in the **ARE** (Eq. 5-6) is identically zero. Although Eq. 5-6 can be solved by the existing algorithms, for example, the Schur method in [L1], a simpler and more efficient algorithm will be presented in the next section for this class of **ARE**.

5.2 A Lyapunov Approach for Obtaining the Inner Denominator

In Section 5.1, the coprime factorization with inner denominator is obtained where an ARE is required to be solved. In this section, a more efficient algorithm will be developed by taking advantage of the block upper triangular structure of the associated Hamiltonian matrix (Eq. 5-3 or 5-9). The computation involved is much simpler compared to that in the general Schur method where the Schur reduction of a $2n \times 2n$ matrix is required [L1]. The idea of invariant subspace corresponding to the eigenvalues in the open *lhp* will be used. Without loss of generality, only the right coprime case is treated here.

Consider the Hamiltonian matrix which arises in the *rcf* with inner denominator,

$$A_H = \begin{bmatrix} A & -BB^T \\ 0 & -A^T \end{bmatrix}$$

where $\mathrm{Re}[\lambda_i(A)] \neq 0$, $i = 1,2,\ldots\ldots,n$ and the pair (A,B) is stabilizable. The algorithm is composed of the following steps:

Step 1

Find the upper Real Schur Form of A_H such that the eigenvalues with negative real parts appears first, i.e.,

$$U^T A U = \begin{bmatrix} A_s & A_{su} \\ 0 & A_u \end{bmatrix}$$

where A_s and $-A_u$ are stable matrices and $U = \begin{bmatrix} U_s & U_u \end{bmatrix}$ is an orthogonal matrix. Note that the columns of U_s are the Schur vectors corresponding to the eigenvalues with negative real parts.

Step 2

Apply the similarity transformation $T_1 = \begin{bmatrix} U^T & 0 \\ 0 & U^T \end{bmatrix}$ to A_H:

$$T_1 A_H T_1^{-1} = \begin{bmatrix} U^T A U & -U^T BB^T U \\ 0 & -U^T A^T U \end{bmatrix} = \begin{bmatrix} A_s & A_{su} & -B_s B_s^T & -B_s B_u^T \\ 0 & A_u & -B_u B_s^T & -B_u B_u^T \\ 0 & 0 & -A_s^T & 0 \\ 0 & 0 & -A_{su}^T & -A_u^T \end{bmatrix}$$

where $B_s = U_s^T B$ and $B_u = U_u^T B$.

Step 3

Solve the Lyapunov equation:

$$A_u Z + Z A_u^T = B_u B_u^T. \tag{5-10}$$

Since the eigenvalues of A_u are in the open rhp, Eq. 5-10 has a unique solution Z. From the assumption of the stabilizability of (A, B), the pair (A_u, B_u) must be controllable. Therefore, $Z > 0$ is concluded from the Lyapunov stability theorem.

Step 4

Define

$$T_2 = \begin{bmatrix} I & 0 & 0 & I \\ 0 & I & 0 & -Z \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Apply similarity transformation T_2 to $T_1 A_H T_1^{-1}$ to get

$$T_2(T_1 A_H T_1^{-1}) T_2^{-1} = \begin{bmatrix} A_s & A_{su} & -A_{su}^T - B_s B_s^T & -A_u^T - A_s + A_{su} Z - B_s B_u^T \\ 0 & A_u & -B_u B_s^T + Z A_{su}^T & 0 \\ 0 & 0 & -A_s^T & 0 \\ 0 & 0 & -A_{su}^T & -A_u^T \end{bmatrix}.$$

Step 5

Form the third similarity transformation T_3 as

$$T_3 = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} .$$

Then

$$T_3(T_2 T_1 A_H T_1^{-1} T_2^{-1})T_3^{-1} = \begin{bmatrix} A_s & -A_u^T - A_s + A_{su} Z - B_s B_u^T & A_{su} & -A_{su}^T - B_s B_s^T \\ 0 & -A_u^T & 0 & -A_{su}^T \\ 0 & 0 & A_u & -B_u B_s^T + Z A_{su}^T \\ 0 & 0 & 0 & -A_s^T \end{bmatrix} . \quad (5\text{-}11)$$

Step 6

Define $T = T_3 T_2 T_1$, then

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1}$$

$$= \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} I & 0 & 0 & I \\ 0 & I & 0 & Z \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \\ 0 & I & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} U \begin{bmatrix} I & I \\ 0 & Z \end{bmatrix} & U \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix} \\ U \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} & U \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \end{bmatrix} .$$

Note that the left upper $n \times n$ matrix in Eq. 5-11 is stable, therefore, the stabilizing solution X is simply

$$X = \left\{ U \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \right\} \left\{ U \begin{bmatrix} I & I \\ 0 & Z \end{bmatrix} \right\}^{-1}$$

$$= \begin{bmatrix} U_s & U_u \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & Z^{-1} \end{bmatrix} \begin{bmatrix} U_s^T \\ U_u^T \end{bmatrix}$$

$$= U_u Z^{-1} U_u^T \quad \geq 0.$$

To summarize the procedure,

(a) Find the upper real Schur form of A such that the eigenvalues with negative real parts appear first, i.e.,

$$U^T A U = \begin{bmatrix} A_s & A_{su} \\ 0 & A_u \end{bmatrix} \quad \text{where} \quad U = \begin{bmatrix} U_s & U_u \end{bmatrix}$$

(b) Solve the Lyapunov equation:

$$A_u Z + Z A_u^T = U_u^T B B^T U_u^T.$$

(c) The stabilizing solution is $X = U_u Z^{-1} U_u^T$.

Remarks

(i)   The method described here is more efficient because the matrix involved in the Schur reduction is only a $n \times n$ in contrast to $2n \times 2n$ in the general Schur method.

(ii)   Since the matrix A_u is already in the quasi-upper triangular form, the solution to Eq. 5-10 can be obtained by simply solving a set of linear equations [B2,G3].

(iii) Z^{-1} can be computed by taking advantage of the symmetry of Z which has order at most n (n is the dimension of A).

(iv)   If the realization of $G(s)$ is already in the balanced form, then the solution $X = diag\left(\dfrac{1}{\sigma_1}, \dfrac{1}{\sigma_{2,}} \ldots\ldots, \dfrac{1}{\sigma_k}\right)$, where $\{\sigma_i\}$ are the second order modes of the unstable subsystem of $G(s)$.

(v)   A special case is when the matrix A is completely unstable (i.e., $Re[\lambda_i(A)] > 0,\ \forall i$) and (A,B) is controllable, then the **ARE**

$$A^T X + XA - XBB^T X = 0$$

is equivalent to the following Lyapunov equation

$$YA^T + AY = BB^T .$$

Since (A,B) is controllable and $Re[\lambda_i(A)] > 0$ guarantee that $Y > 0$, $X = Y^{-1}$ is the unique positive definite stabilizing solution of the **ARE**.

5.3 Spectral Factorizations

In this section, without loss of generality, $G(s)$ is assumed to be stable since any $G \in RL_\infty$ may be factored using Corollary 5.1.2 to obtain a stable numerator $\tilde{N}$ such that $\tilde{N}^*\tilde{N} = G^*G$. The objective now is to derive a state-space formula for the spectral factor of $(\gamma^2 I - G^*G)$ (or $(\gamma^2 I - GG^*)$).

Theorem 1 (Spectral Factorization)

Assume $G(s) \in RH_\infty^{t \times m}$ and $\gamma > ||G(s)||_\infty$. Then, there exists a $M \in RH_\infty^{m \times m}$ with stable inverse such that $M^*M = \gamma^2 I - G^*G$ with

$$M = \left[\begin{array}{c|c} A & B \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array}\right] \tag{5-12}$$

where

$$R_D = \gamma^2 I - D^T D > 0$$

$$K_c = -R_D^{-1}(B^T X - D^T C)$$

and

$$X = \mathbf{Ric} \left[\begin{array}{cc} A + BR_D^{-1}D^T C & -BR_D^{-1}B^T \\ C^T(I + DR_D^{-1}D^T)C & -(A + BR_D^{-1}D^T C)^T \end{array}\right].$$

Proof

Let

$$\Gamma = \gamma^2 I - G^*G = \left[\begin{array}{cc} B^T(-sI - A^T)^{-1} & I \end{array}\right] \left[\begin{array}{cc} -C^T C & -C^T D \\ -D^T C & R_D \end{array}\right] \left[\begin{array}{c} (sI - A)^{-1}B \\ I \end{array}\right].$$

Since $\gamma > ||G||_\infty$, $\Gamma(j\omega) > 0$. The minimality of the realization of G guarantees that (A, B) is controllable and $(-C^T C, A)$ is observable. Thus, from Corollary 2.3.2, there exists $M \in RH_\infty^{m \times m}$ and $M^{-1} \in RH_\infty$ such that $\Gamma = M^*M$ and a particular realization is

$$M = \left[\begin{array}{c|c} A & B \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array}\right]$$

where

$$K_c \;=\; -R_D^{-1}(B^T X - D^T C)$$

and

$$X \;=\; \mathbf{Ric}\begin{bmatrix} A + BR_D^{-1}D^T C & -BR_D^{-1}B^T \\ C^T(I + DR_D^{-1}D^T)C & -(A + BR_D^{-1}D^T C)^T \end{bmatrix}.$$

Since G is stable, it is concluded that $M \in RH_\infty^{m \times m}$.

<div align="right">QED</div>

The following corollary is the dual result of Theorem 1.

Corollary 2

With the same assumptions as in Theorem 1, then there exists a $\tilde{M} \in RH_\infty^{p \times p}$ with stable inverse such that $\tilde{M}\tilde{M}^* = \gamma^2 I - GG^*$ with

$$\tilde{M} \;=\; \left[\begin{array}{c|c} A & -K_f \tilde{R}_D^{1/2} \\ \hline C & \tilde{R}_D^{1/2} \end{array}\right] \tag{5-13}$$

where

$$\tilde{R}_D \;=\; \gamma^2 I - DD^T \;>\; 0$$

$$K_f \;=\; -(YC^T - B^T D)\tilde{R}_D^{-1}$$

and

$$Y \;=\; \mathbf{Ric}\begin{bmatrix} (A + BD^T\tilde{R}_D^{-1}C)^T & -C^T\tilde{R}_D^{-1}C \\ B(I + D^T\tilde{R}_D^{-1}D)B^T & -(A + BD^T\tilde{R}_D^{-1}C) \end{bmatrix}.$$

Remark

The notation $(\gamma^2 I - G^*G)^{1/2}$ ($(\gamma^2 I - GG^*)^{1/2}$) will be used to denote the spectral factor M ($\tilde{M}$) in Theorem 1 (Corollary 2).

Theorem 3

If $\gamma = ||G(s)||_\infty$ but $\gamma > \bar{\sigma}[G(\infty)]$ in Theorem 1, then there exists a $M \in RH_\infty^{m \times m}$ with M^{-1} analytic in the open rhp such that $M^*M = \gamma^2 I - G^*G$ with

$$M = \left[\begin{array}{c|c} A & B \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array} \right]$$

where

$$R_D = \gamma^2 I - D^T D > 0$$

$$K_c = -R_D^{-1}(B^T X - D^T C)$$

and X is the unique solution to the **ARE**

$$(A + BR_D^{-1}D^T C)^T X + X(A + BR_D^{-1}D^T C) - XBR_D^{-1}B^T X + C^T(I + DR_D^{-1}D^T)C = 0 \quad (5\text{-}14)$$

such that $\text{Re}[\lambda_i (A + BK_c)] \leq 0, \ \forall i$.

Proof

Let

$$\Gamma = \gamma^2 I - G^*G = \left[B^T(-sI - A^T)^{-1} \ I \right] \begin{bmatrix} -C^T C & -C^T D \\ -D^T C & R_D \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}$$

By assumption, $\Gamma(j\omega) \geq 0$ and $R_D = \Gamma(\infty) > 0$. Since (A,B) is controllable, from a theorem of J.C. Willems [W1], there exists a unique solution X to Eq. 5-14 such that

$$A + BR_D^{-1}D^T C - BR_D^{-1}B^T X = A + BK_c$$

has no eigenvalues in the open rhp .

QED

Corollary 2 can be generalized in a similar way to the case where $\gamma = ||G||_\infty$ but $\gamma > \bar{\sigma}[G(\infty)]$ such that $\bar{M}^{-1}$ is analytic in the open rhp .

5.4 Balanced Realizations and Best Approximation Problems

In Section 3.2, it was shown that there exists an optimal solution $Q \in H_\infty$ to Eq. 3-14 such that $||R - Q||_\infty = ||H_R||$. The same equation will be reconsidered in this section for real-rational R. The idea of balanced realization will play a key role in this problem.

Recall that a stable transfer matrix $G(s)$ with minimal realization $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is called "balanced" if there exists a diagonal matrix

$$\Sigma = diag\ (\sigma_1, \sigma_2, \ldots, \sigma_n) \qquad \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0,$$

such that the following two Lyapunov equations are satisfied:

$$A\Sigma + \Sigma A^T = -BB^T$$

and

$$A^T\Sigma + \Sigma A = -C^T C.$$

where $\{\sigma_i\}$ are called the "2nd order modes" of the system [M4,P2].

If $G(s)$ is completely unstable (i.e., no eigenvalues in the closed lhp), then $G(s)$ is said to be balanced if $G(-s)$ is. The method for obtaining balanced realization is not unique which can be found in [E1,L2,M4]. Note that balanced realization is also very useful in model-reduction [E1,G2,K7].

The following lemma shows the relation between the norm of a Hankel operator and the balanced realization of a completely unstable transfer function matrix G.

Lemma 1

Assume G is real-rational and completely unstable in Eq. 3-14. Then

$$||H_G|| = \sigma_1$$

where σ_1 is the largest "2nd order mode" of G.

This lemma is very useful, since it relates that the Hankel-norm can be computed easily using state-space methods. In fact, even if the realization is not balanced, the norm is equal to $\lambda_{max}^{1/2}(XY)$ where X and Y are the controllability and observability gramians respectively and can be computed by solving the corresponding Lyapunov equations:

$$AX + XA^T = BB^T$$

and

$$A^TY + YA = C^TC.$$

Now, consider the best approximation for the real-rational case:

$$\gamma_o = \min_{Q \in RH_\infty} || R - Q ||_\infty , \quad R(-s) \in RH_\infty . \tag{5-15}$$

The question is: does Eq. 5-15 have a solution which is also real-rational? The answer is yes and the proof can be found in [A1,A2,B3,D16,G2].

Theorem 2

If R is real-rational in Eq. 5-15, then there exists a best approximation $Q \in RH_\infty$.

Theorem 3

Assume that Q_{opt} is the best approximation in Eq. 5-15, then

(i)   if $p = 1$ or $m = 1$, Q_{opt} is unique and $\frac{1}{\gamma_o}(G - Q_{opt})$ is all-pass.

(ii)   if $p \neq 1$ and $m \neq 1$, then Q_{opt} is not unique and

$$\frac{1}{\gamma_o}\overline{\sigma}(G - Q_{opt})(j\omega) = 1 \quad \forall \omega \in \mathbf{R} .$$

There are various algorithms available to solve this problem [B1,B3,D16,G2]. However, the Glover's algorithm [G2] appears to be the most efficient one so far and will be reviewed next.

Without loss of generality, R is assumed to be strictly proper, completely unstable (the stable part of R can always be absorbed into Q) and square (if not square, add rows or columns of zeros). Let $R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ be a minimal realization. Then the optimal $Q \in RH_\infty$ can be constructed using the following steps:

Step 1

Find a balanced realization of R.

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right].$$

Thus both the controllability and observability gramians are diagonal and equal to

$$P = \begin{bmatrix} \sigma I, & 0 \\ 0 & \Sigma \end{bmatrix}$$

where $\sigma > ||\Sigma||$, i.e., $r =$ multiplicity of σ. Partition A, B, and C accordingly as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Step 2

Choose $\hat{D}$ such that

$$\hat{D} B_1^T + \sigma C_1 = 0 \tag{5-16}$$

$$\hat{D} \hat{D}^T = \sigma^2 I. \tag{5-17}$$

Step 3

Set

$$\hat{B} = -(\sigma^2 I - \Sigma^2)^{-1}(\Sigma B_2 + \sigma C_2^T \hat{D}) \tag{5-18}$$

$$\hat{A} = (-A_{22} + B_2 \hat{B}^T)^T \tag{5-19}$$

and

$$\hat{C} = C_2 \Sigma + \hat{D} B_2^T \ . \tag{5-20}$$

Step 4

Let

$$Q = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array} \right]$$

Then, Q is an optimal solution to Eq. 5-15 [G2].

5.5 State-Space Solutions of General Distance Problems

Combining the results from previous sections of this chapter, in this section, the state-space realization of the optimal solution of the general distance problem will be derived step-by-step. Although the formula may not look very simple in its appearance, the implementation is quite straightforward. Once again, the 2-block problem will be considered first, and the results are then generalized to the 4-block problem.

(A) State-space optimal solutions of the 2-block GDP

Without loss of generality, both R_1 and R_2 are assumed to be completely unstable with the realizations

$$R_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \quad \text{and} \quad R_2 = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

where the realization R_2 is minimal. The solution can be found from the following steps:

Step 1 (Left coprime factorization with inner denominator) [Corollary 5.1.2]

$$R_2 = \tilde{M}^{-1}\tilde{N} \quad \text{with} \quad \tilde{N} = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline C & D \end{array}\right] \quad \in RH_\infty$$

where $\tilde{A} = A + HC$, $\tilde{B} = B + HD$ and H is the observer gain such that $\tilde{A}$ is asymptotically stable.

Step 2 (Spectral factorization of $(\gamma^2 I - R_2^* R_2)$) [Theorem 5.3.1]

Find the spectral factor M such that $M^* M = \gamma^2 I - \tilde{N}^* \tilde{N} = \gamma^2 I - R_2^* R_2$.

$$M = \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array} \right]$$

where

$$R_D = \gamma^2 I - D^T D$$
$$K_c = -R_D^{-1}(\bar{B}^T X - D^T C)$$

and

$$X = \mathbf{Ric} \left[\begin{array}{cc} \bar{A} + \bar{B}R_D^{-1}D^T C & -\bar{B}R_D^{-1}\bar{B}^T \\ C^T(I + DR_D^{-1}D^T)C & -(\bar{A} + \bar{B}R_D^{-1}D^T C)^T \end{array} \right]$$

where $\mathbf{Ric}\,(A_H)$ denotes the stabilizing solution of the Riccati equation with the associated Hamiltonian matrix A_H.

Step 3

Form

$$G = R_1 M^{-1}$$
$$= \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array} \right] * \left[\begin{array}{c|c} \bar{A} + \bar{B}K_c & \bar{B}R_D^{-1/2} \\ \hline K_c & R_D^{-1/2} \end{array} \right]$$
$$= \left[\begin{array}{cc|c} A_1 & B_1 K_c & B_1 R_D^{-1/2} \\ 0 & \bar{A} + \bar{B}K_c & \bar{B}R_D^{-1/2} \\ \hline C_1 & D_1 K_c & D_1 R_D^{-1/2} \end{array} \right] . \tag{5-21}$$

Step 4 (Spectral Decomposition)

(i) Solve the following Sylvester equation for Z:

$$-A_1 Z + Z(\bar{A} + \bar{B}K_c) = -B_1 K_c . \tag{5-22}$$

Since both $-(A_1)$ and $(\bar{A} + \bar{B}K_c)$ are stable, Eq. 5-22 has a unique solution.

(ii) Conjugating the states in Eq. 5-21 by $T = \left[\begin{array}{cc} I & Z \\ 0 & I \end{array} \right]$,

$$G = \begin{bmatrix} A_1 & 0 & (B_1 + Z\bar{B})R_D^{-1/2} \\ 0 & \bar{A} + \bar{B}K_c & \bar{B}R_D^{-1/2} \\ \hline C_1 & D_1K_c - C_1Z & D_1R_D^{-1/2} \end{bmatrix}$$

$$= \begin{bmatrix} A_1 & (B_1 + Z\bar{B})R_D^{-1/2} \\ \hline C_1 & 0 \end{bmatrix} + \begin{bmatrix} \bar{A} + \bar{B}K_c & \bar{B}R_D^{-1/2} \\ \hline D_1K_c - C_1Z & D_1R_D^{-1/2} \end{bmatrix}$$

$$= G_u + G_s$$

where G_s is stable and G_u is completely unstable.

Step 5 (Computation of the Hankel-norm)

Compute the Hankel-norm of $G_u(s)$, called $f(\gamma)$, which is equal to $[\lambda_{max}(W_c W_s)]^{1/2}$ where W_c and W_s are the controllability and observability gramians of $G_u(-s)$ respectively, i.e., W_c and W_s are the solutions of the following two Lyapunov equations

$$A_1 W_c + W_c A_1^T = (B_1 + Z\bar{B})R_D^{-1}(B_1 + Z\bar{B})^T$$

and

$$A_1^T W_s + W_s A_1 = C_1^T C_1$$

respectively.

(a)   if $f(\gamma) > 1$, it means that the value of γ is too small. A larger γ should be chosen, and the process will restart from **Step 2** .

(b)   if $f(\gamma) = 1$, from Theorem 4.2.8, the corresponding γ is the minimal achievable norm in (2gdp).

(c)   if $f(\gamma) < 1$, From Theorems 4.2.5 and 4.2.9 two cases can occur:

(i) $\gamma_s = \gamma$.

and

(ii) $\gamma_s < \gamma$.

Case (i) can happen only when $\gamma_s = ||R_2||_\infty$. It can always be detected by letting $\gamma = ||R_2||_\infty + \epsilon_s$ in the beginning of the iterative process where ϵ_s is a very small but positive real number, for example, 10^{-6}. If the corresponding $f(\gamma) < 1$, then γ_s must be equal to $||R_2||_\infty$. This special case will be discussed later in this section. Assume that (ii) is the case, it means that the value of γ is over-estimated. The value of γ should be decreased. Then the process will restart from **Step 2** .

Step 6 (Best Approximation) [Section 5.4]

If $f(\gamma) = 1$, Glover's algorithm which was reviewed in Section 4 can be used to find the best approximation $\tilde{Q}_s \in RH_\infty$ of G_u. Assume $\tilde{Q}_s$ has the following realization:

$$\tilde{Q}_s = \left[\begin{array}{cc} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right] .$$

Step 7

(i) Form

$$G_s M = \left[\begin{array}{c|c} \tilde{A} + \tilde{B}K_c & \tilde{B}R_D^{-1/2} \\ \hline D_1 K_c - C_1 Z & D_1 R_D^{-1/2} \end{array}\right] * \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array}\right]$$

$$= \left[\begin{array}{cc|c} \tilde{A} + \tilde{B}K_c & -\tilde{B}K_c & \tilde{B} \\ 0 & \tilde{A} & \tilde{B} \\ \hline D_1 K_c - C_1 Z & -D_1 K_c & D_1 \end{array}\right] . \tag{5-23}$$

(ii) Conjugating the states in Eq. 5-23 by $T = \left[\begin{array}{cc} I & -I \\ 0 & I \end{array}\right]$ then

$$G_s M = \left[\begin{array}{cc|c} \tilde{A} + \tilde{B}K_c & 0 & 0 \\ 0 & \tilde{A} & \tilde{B} \\ \hline D_1 K_c - C_1 Z & -C_1 Z & D_1 \end{array}\right] = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline -C_1 Z & D_1 \end{array}\right] . \tag{5-24}$$

Step 8

Form

$$\tilde{Q}_{\,,} M \;=\; \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right] \;*\; \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline -R_D^{1/2}K_c & R_D^{1/2} \end{array}\right]$$

$$=\; \left[\begin{array}{cc|c} A_Q & -B_Q R_D^{1/2}K_c & B_Q R_D^{1/2} \\ 0 & \tilde{A} & \tilde{B} \\ \hline C_Q & -D_Q R_D^{1/2}K_c & D_Q R_D^{1/2} \end{array}\right].$$

Note that Eq. 5-24 can be written as

$$G_{\,,} M \;=\; \left[\begin{array}{cc|c} A_Q & -B_Q R_D^{1/2}K_c & B_Q R_D^{1/2} \\ 0 & \tilde{A} & \tilde{B} \\ \hline C_Q & -C_1 Z & D_1 \end{array}\right].$$

Therefore,

$$Q_{\,,} \;=\; (\tilde{Q}_{\,,}\; +\; G_{\,,})M \;=\; \left[\begin{array}{cc|c} A_Q & -B_Q R_D^{1/2}K_c & B_Q R_D^{1/2} \\ 0 & \tilde{A} & \tilde{B} \\ \hline C_Q & -D_Q R_D^{1/2}K_c - C_1 Z & D_Q R_D^{1/2} + D_1 \end{array}\right]. \qquad (5\text{-}25)$$

If $\gamma_{\,,} = ||R_2||_\infty$ and $\gamma_{\,,} > \bar{\sigma}[R_2(\infty)]$, then **Step 1-8** can still be used to find the optimal $Q_{\,,}$, the only difference is in **Step 2** where the factor M will have zeros on the $j\omega$-axis, i.e., $\tilde{A} + \tilde{B}K_c$ has no eigenvalues in the open rhp. Since $(\tilde{A} + \tilde{B}K_c)$ does not appear in Eq. 5-25, the corresponding $Q_{\,,}$ is still in RH_∞. If $\gamma_{\,,} = ||R_2||_\infty$ but $\gamma_{\,,} = \bar{\sigma}[R_2(\infty)]$, then M does not have full rank at ∞, which is equivalent to saying that M has no inverse in RL_∞. Therefore, the γ-iteration doesn't apply to this particular problem. Although $R_2(\infty)$ can always be perturbed by some $\epsilon > 0$ and use the method described here (**Step 1-8**), there is no theoretical support for this argument. Further research needs to be done in this special case. Theorem 3.2.2 is therefore proved except the last case.

(B) State-space optimal solutions of the 4-block GDP

The 4-block case is much more complicated than 2-block case. However, a closed form solution can still be found. Without loss of generality, R is assumed to be completely unstable (Theorem 2.3.6). Due to the complexity of the problem, the realization in the first two steps will not be shown which are not critical in the process.

The solution can be found from the following steps:

Step 1

(i) Find factors $S = (\gamma^2 I - R_{22}^* R_{22})^{1/2}$ and $\tilde{S} = (\gamma^2 I - R_{22} R_{22}^*)^{1/2}$.

Note that S and $\tilde{S}$ need not be spectral factors. See the Remark (ii) after Theorem 4.1.3.

(ii) Form $L = R_{12} S^{-1}$ and $\tilde{L} = \tilde{S}^{-1} R_{21}$.

Step 2

(i) Find the rcf with inner denominator for L,

$$L = N_L M_L^{-1} .$$

(ii) Find the lcf with inner denominator for $\tilde{L}$,

$$\tilde{L} = M_{\tilde{L}}^{-1} N_{\tilde{L}} .$$

Assume that N_L and $N_{\tilde{L}}$ have the following minimal realizations:

$$N_L = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array} \right] \quad \text{and} \quad N_{\tilde{L}} = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array} \right] .$$

Step 3 (Spectral factorization)

(i) Using Theorem 4.1.1, find the spectral factor M such that

$$MM^* = I - N_L N_L^* = I - LL^* .$$

where

$$M = \left[\begin{array}{cc} A_1 & -K_f \bar{R}_D^{1/2} \\ \hline C_1 & \bar{R}_D^{1/2} \end{array}\right] .$$

(ii) Find the spectral factor $\tilde{M}$ such that

$$\tilde{M}^* \tilde{M} = I - N_{\tilde{L}}^* N_{\tilde{L}} = I - \tilde{L}^* \tilde{L}$$

where

$$\tilde{M} = \left[\begin{array}{cc} A_2 & B_2 \\ \hline -R_D^{1/2} K_c & R_D^{1/2} \end{array}\right]$$

For simplicity, the expressions for K_c, K_f, R_D and $\tilde{R}_D$ will not be written down explicitly.

Step 4

(i) Form the linear fractional transformation

$$\mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^*) = \frac{1}{\gamma}\left\{ R_{11} + R_{12}(\gamma^2 I - R_{22}^* R_{22})^{-1} R_{22}^* R_{21} \right\} .$$

Assume that

$$\mathbf{F}_l(\frac{1}{\gamma}R, \frac{1}{\gamma}R_{22}^*) = G_s + G_u = \left[\begin{array}{cc} A_s & B_s \\ C_s & D_s \end{array}\right] + \left[\begin{array}{cc} A_u & B_u \\ C_u & 0 \end{array}\right]$$

where G_s is stable and G_u is completely unstable.

Step 5

Form the product

$$\begin{aligned}
\bar{R} &= M^{-1} G_u \tilde{M}^{-1} \\
&= \left[\begin{array}{cc|c} A_1 + K_f C_1 & K_f \\ \hline \bar{R}_D^{-1/2} C_1 & \bar{R}_D^{-1/2} \end{array}\right] * \left[\begin{array}{c|c} A_u & B_u \\ \hline C_u & 0 \end{array}\right] * \left[\begin{array}{c|c} A_2 + B_2 K_c & B_2 R_D^{-1/2} \\ \hline K_c & R_D^{-1/2} \end{array}\right] .
\end{aligned}$$

Step 6 (Spectral decomposition)

$$\bar{R} = \bar{R}_s + \bar{R}_u$$

with

$$\bar{R}_s = \left[\begin{array}{cc|c} A_1 + K_f\,C_1 & -Z_1 B_u K_c & -Z_1 B_u R_D^{-1/2} \\ 0 & A_2 + B_2 K_c & B_2 R_D^{-1/2} \\ \hline \bar{R}_D^{-1/2} C_1 & \bar{R}_D^{-1/2}(C_1 Z_1 + C_u)Z_2 & 0 \end{array}\right]$$

and

$$\bar{R}_u = \left[\begin{array}{c|c} A_u & (B_u - Z_2 B_2)R_D^{-1/2} \\ \hline \bar{R}_D^{-1/2}(C_1 Z_1 + C_u) & 0 \end{array}\right]$$

where

(i) Z_1 is the unique solution of the following Sylvester equation:

$$(A_1 + K_f\,C_1)Z_1 - Z_1 A_u + K_f\,C_u = 0 .$$

(ii) Z_2 is the unique solution of the following Sylvester equation:

$$A_u Z_2 - Z_2(A_2 + B_2 K_c) + B_u K_c = 0 .$$

Step 7 (Best approximation)

Find the best approximation of $\bar{R}_u$, i.e.,

$$f(\gamma) = \min_{\tilde{Q}\,\in\,RH_\infty} ||\,\bar{R}_u - \tilde{Q}\,||_\infty . \tag{5-26}$$

Using Glover's algorithm, $f(\gamma)$ can be computed before the best approximation of G_u is found. If $f(\gamma) > 1$, it means that the value of γ is too small. A larger γ should be chosen, and the process will restart from **Step 1** .

If $f(\gamma)$ is much smaller than 1, in general, this implies that the value of γ is over-estimated and a smaller value for γ should be used. It is desired that $\bar{\gamma}_s$ is equal or as close as possible to 1. However, this is not true for every case.

If $f(\gamma) = 1$, then the optimal solution to Eq. 5-23 is $\tilde{Q}$, with the following realization:

$$\tilde{Q}_. = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right].$$

Therefore, $\overline{Q}_. = \tilde{Q}_. + \overline{R}_.$ is the best approximation of $\overline{R}$. Correspondingly,

$$Q_. = M(\overline{Q}_. + M^{-1}G_. \tilde{M}^{-1})\tilde{M} = M\overline{Q}_. \tilde{M} + G_. .$$

Step 8

(i) Form

$$\overline{Q}_. = \tilde{Q}_. + \overline{R}_.$$

$$= \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right] + \left[\begin{array}{cc|c} A_1 + K_f C_1 & -Z_1 B_u K_c & -Z_1 B_u R_D^{-1/2} \\ 0 & A_2 + B_2 K_c & B_2 R_D^{-1/2} \\ \hline \tilde{R}_D^{-1/2} C_1 & \tilde{R}_D^{-1/2}(C_1 Z_1 + C_u)Z_2 & 0 \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} A_1 + K_f C_1 & 0 & -Z_1 B_u K_c & -Z_1 B_u R_D^{-1/2} \\ 0 & A_Q & 0 & B_Q \\ 0 & 0 & A_2 + B_2 K_c & B_2 R_D^{-1/2} \\ \hline \tilde{R}_D^{-1/2} C_1 & C_Q & \tilde{R}_D^{-1/2}(C_1 Z_1 + C_u)Z_2 & D_Q \end{array}\right].$$

(ii) Form

$$M\overline{Q}_. \tilde{M}$$

$$= \left[\begin{array}{c|c} A_1 & -K_f \tilde{R}_D^{1/2} \\ \hline C_1 & \tilde{R}_D^{1/2} \end{array}\right] * \left[\begin{array}{ccc|c} A_1 + K_f C_1 & 0 & -Z_1 B_u K_c & -Z_1 B_u R_D^{-1/2} \\ 0 & A_Q & 0 & B_Q \\ 0 & 0 & A_2 + B_2 K_c & B_2 R_D^{-1/2} \\ \hline \tilde{R}_D^{-1/2} C_1 & C_Q & \tilde{R}_D^{-1/2}(C_1 Z_1 + C_u)Z_2 & D_Q \end{array}\right] * \left[\begin{array}{c|c} A_2 & B_2 \\ \hline -R_D^{1/2} K_c & R_D^{1/2} \end{array}\right]$$

$$= \left[\begin{array}{ccccc|c} A_1 & -K_f C_1 & K_f \tilde{R}_D^{1/2} C_Q & -K_f(C_1 Z_1 + C_u)Z_2 & K_f \tilde{R}_D^{1/2} D_Q R_D^{1/2} K_c & -K_f \tilde{R}_D^{1/2} D_Q R_D^{1/2} \\ 0 & A_1 + K_f C_1 & 0 & -Z_1 B_u K_c & Z_1 B_u K_c & -Z_1 B_u \\ 0 & 0 & A_Q & 0 & -B_Q R_D^{1/2} K_c & B_Q R_D^{1/2} \\ 0 & 0 & 0 & A_2 + B_2 K_c & -B_2 K_c & B_2 \\ 0 & 0 & 0 & 0 & A_2 & B_2 \\ \hline C_1 & C_1 & \tilde{R}_D^{1/2} C_Q & (C_1 Z_1 + C_u)Z_2 & \tilde{R}_D^{1/2} D_Q R_D^{1/2} K_c & \tilde{R}_D^{1/2} D_Q R_D^{1/2} \end{array}\right].$$

Step 9

Conjugating the states in the above realization by

$$T = \begin{bmatrix} I & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & -I \\ 0 & 0 & 0 & 0 & I \end{bmatrix}.$$

Then

$M\bar{Q}_s\bar{M}$

$$= \begin{bmatrix} A_1 & 0 & -K_f\bar{R}_D^{1/2}C_Q & -K_f(C_1Z_1+C_u)Z_2-Z_1B_uK_c & -K_f(C_1Z_1+C_u)Z_2-K_fD_QK_c & -K_fD_Q-Z_1B \\ 0 & A_1+K_fC_1 & 0 & -Z_1B_uK_c & 0 & -Z_1B_u \\ 0 & 0 & A_Q & 0 & -B_QR_D^{1/2}K_c & B_QR_D^{1/2} \\ 0 & 0 & 0 & A_2+B_2K_c & 0 & 0 \\ 0 & 0 & 0 & 0 & A_2 & B_2 \\ \hline C_1 & 0 & \bar{R}_D^{1/2}C_Q & (C_1Z_1+C_u)Z_2 & (C_1Z_1+C_u)Z_2-\bar{R}_D^{1/2}D_QR_D^{1/2}K_c & \bar{R}_D^{1/2}D_QR_D^{1/2} \end{bmatrix}$$

$$= \begin{bmatrix} A_1 & -K_f\bar{R}_D^{1/2}C_Q & -K_f(C_1Z_1+C_u)Z_2-K_fD_QK_c & -K_fD_Q-Z_1B_u \\ 0 & A_Q & -B_QR_D^{1/2}K_c & B_QR_D^{1/2} \\ 0 & 0 & A_2 & B_2 \\ \hline C_1 & \bar{R}_D^{1/2}C_Q & (C_1Z_1+C_u)Z_2-\bar{R}_D^{1/2}D_QR_D^{1/2}K_c & \bar{R}_D^{1/2}D_QR_D^{1/2} \end{bmatrix}$$

Therefore,

$Q_s = M\bar{Q}_s\bar{M} + G_s$

$$= \begin{bmatrix} A_s & 0 & 0 & 0 & B_s \\ 0 & A_1 & K_f\bar{R}_D^{1/2}C_Q & -K_f(C_1Z_1+C_u)Z_2-K_fD_QK_c & -K_fD_Q-Z_1B_u \\ 0 & 0 & A_Q & -B_QR_D^{1/2}K_c & B_QR_D^{1/2} \\ 0 & 0 & 0 & A_2 & B_2 \\ \hline C_s & C_1 & \bar{R}_D^{1/2}C_Q & (C_1Z_1+C_u)Z_2-\bar{R}_D^{1/2}D_QR_D^{1/2}K_c & D_s+\bar{R}_D^{1/2}D_QR_D^{1/2} \end{bmatrix}.$$

CHAPTER 6
NUMERICAL ASPECTS OF
H_2 AND H_∞ OPTIMIZATION PROBLEMS

In this chapter, the H_2 and H_∞ optimal control problems are summarized using a unified general framework. In addition, numerical aspects of the algorithms and the issue of model-reduction are also discussed.

Section 6.1 recapitulates the key steps required to solve the H_2 and H_∞ optimal control problems for finite-dimensional linear time-invariant systems. One of the major contributions of this work is the development of a complete state-space approach to obtain the optimal solution. The implementation of the algorithms is straightforward involving only standard matrix operations and linear algebra techniques. In Section 6.2, some numerical aspects of the algorithms are discussed. Since the H_∞ synthesis results in a high-order optimal controller, and because of computational and other practical limitations, it is desirable that some form of model-reduction be used. In particular, it is shown in Section 6.3 that model-reduction can be used in the H_∞ synthesis procedure to obtain a lower order (suboptimal) controller with a priori bounds on the degree of suboptimality.

6.1 Summary of H_2 and H_∞ Optimal Control

This section is a recapitulation of the results developed in the previous chapters to give a complete algorithm for solving H_2 and H_∞ optimal control problems. The key steps are summarized as follows:

Step 1: Parametrization

Find J so that the substitution $K = \mathbf{F}_l(J,Q)$ yields

$$
\begin{aligned}
\mathbf{F}_l(P,K) &= \mathbf{F}_l(P,\mathbf{F}_l(J,Q)) \\
&= \mathbf{F}_l(T,Q) \\
&= T_{11} + T_{12}QT_{21}
\end{aligned}
$$

with the additional requirement that $T \in RH_\infty$ and

$$\mathbf{F}_l(P,K) \quad \text{internally stable}$$

$$\text{if and only if} \quad Q \in RH_\infty.$$

This parametrizes all stabilizing Ks in terms of a stable $Q \in H_\infty$ in addition to providing an affine parametrization of all (internally) stable $\mathbf{F}_l(P,K)$. This "Youla parametrization" [D5,Y2] is developed in Section 2.2. In particular, explicit formula for J and T_{ij}'s are derived using the observer-based stabilization method (Eqs. 2-33 and 2-34).

Step 1a: Coprime Factorization with Inner Numerator

A further requirement is that T_{12} and T_{21} are inner and co-inner with the nonsingular constant matrix multiple respectively. Methods for obtaining the particular parametrizations which achieve this are developed in Section 2.3. By appropriate scaling, the new affine parametrization of the closed-loop transfer matrix becomes

$$T_{11} - N_{12}\hat{Q}N_{21}$$

where N_{12} is inner and N_{21} is co-inner with realizations shown in Eqs. 2-46 through 2-49.

Step 2: Unitary Invariance

Find the CIFs $N_\perp$ and $\bar{N}_\perp$ so that $\begin{bmatrix} N_{12} & N_\perp \end{bmatrix}$ and $\begin{bmatrix} N_{21} \\ \tilde{N}_\perp \end{bmatrix}$ are *square* and inner (also Section 2.3 (Eqs. 2-46 through 2-49)). Then pre- and post-multiply by $\begin{bmatrix} N_{12} & N_\perp \end{bmatrix}^*$ and $\begin{bmatrix} N_{21} \\ \tilde{N}_\perp \end{bmatrix}^*$ to yield (Eq. 2-52)

$$\left| \left| T_{11} - N_{12} \hat{Q} N_{21} \right| \right|_\alpha = \left| \left| \begin{bmatrix} R_{11} - \hat{Q} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right| \right|_\alpha \qquad (\alpha = 2, \infty)$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} T_{11} \begin{bmatrix} N_{21}^* & \tilde{N}_\perp^* \end{bmatrix}$$

The solutions to the H_2 and H_∞ optimization differ completely in the following step.

Step 3: Projection/Dilation

(A) H_2 optimization:

For $\alpha = 2$, the unique optimal solution for $\hat{Q}$ is immediate from Eq. 2-54:

$$\hat{Q}_{opt} = \mathbf{P}_{H_2}(R_{11})$$

(B) H_∞ optimization:

For $\alpha = \infty$, the general distance problem is solved using γ-iteration (Chapters 3, 4 and 5). The iterative procedure is the following:

(i)   compute the lower and upper bounds (Section 4.1).

(ii)   reduce the GDP to the best approximation (Sections 4.1 and 5.3).

(iii)   find the Hankel norm of the corresponding approximation problem (Section 5.4).

Iterate on (ii) an (iii) using the properties of γ-iteration (continuity, monotonicity and convexity) until the optimal γ (or arbitrarily closed to) is found (Section 4.2).

(iv)   derive the state-space formula for $\hat{Q}_{opt}$ (Sections 5.4 and 5.5).

The complete state-space procedure for both 2- and 4-block GDPs are shown in Section 5.5.

Step 4: Recovery of the optimal controller

(i) First, recover Q_{opt} from $\hat{Q}_{opt}$ through Eq. 2-51:

$$Q_{opt} = -R_D^{-1/2}\hat{Q}_{opt}\bar{R}_D^{-1/2}$$

(ii) The optimal controller K_{opt} can be recovered easily from the LFT

$$K_{opt} = \mathbf{F}_l(J, Q_{opt})$$

where J is shown as in Eq. 2-33.

The above steps are also summarized in Figure 6-1.

$$\min_{K \in R_p} \left\{ \| F_l(P,K) \|_\alpha \mid \text{internal stability} \right\}$$

Parametrization
$$K = F_l(J,Q)$$

$$\min_{Q \in H_\infty} \| T_{11} + T_{12}QT_{21} \|_\alpha$$

$$\min_{Q \in H_\infty} \| T_{11} - N_{12}QN_{21} \|_\alpha \qquad \text{where} \qquad \begin{array}{l} N_{12}^* N_{12} = I \\ N_{21}N_{21}^* = I \end{array}$$

$$\min_{Q \in H_\infty} \left\| \begin{bmatrix} R_{11} - \hat{Q} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\alpha$$

$\alpha = 2$
Projection

$\alpha = \infty$
GDP

$$\hat{Q}_{opt} , \ Q_{opt}$$
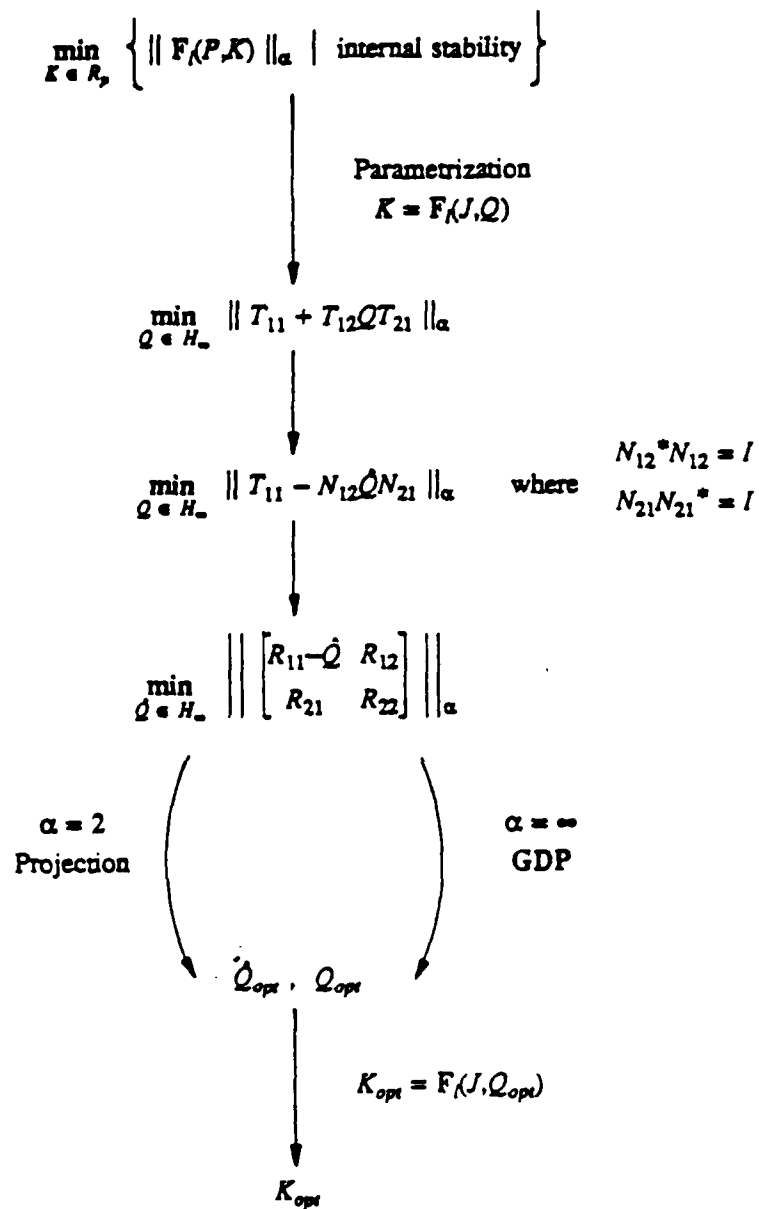
$$K_{opt} = F_l(J, Q_{opt})$$

$$K_{opt}$$

Figure 6-1.

6.2 Numerical Considerations

The implementation of the procedure described in Section 6.1 is quite straightforward since at each step, the state-space formula are also derived in parallel to the theory. Therefore, only standard matrix operations (addition, subtraction, multiplication, inversion) and the familiar linear algebra techniques (singular value decomposition, eigenvalue/eigenvector decomposition, Schur decomposition, etc.) are required essentially. They constitute the nucleus of the entire approach and there are very reliable software packages available [D6,G1] for this purpose.

Most of algorithms developed require solving some particular algebraic equation at an intermediate stage. They are ARE, Lyapunov, and Sylvester equations. Therefore, robust algorithms for solving these equations are extremely critical to a successful implementation. Many reliable algorithms can be found in the literature, (for example, [B2,G3,H1,L1,V1]).

Note that the minimality (or stabilizability/detectability) condition is an important assumption in most of algorithms (coprime factorization, spectral factorization, etc.). To guarantee this condition, a simple procedure is recommended here which is closely related to the notion of "balanced" realization (Section 5.4).

Recall that, for a given stable transfer matrix $G = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, if the realization is balanced (Section 5.4), it must be minimal. Therefore, minimality can be achieved by "balancing". Although all the balancing algorithms developed so far require the given realization (before balancing) to be minimal [L2,M4], they can be modified to handle the non-minimal case by removing the singularity of controllability and/or observability gramians. The (balanced) minimal realization can be obtained similarly for the completely unstable system (see Section 5.4). If the given realization of G has both stable and unstable eigenvalues (but not on the $j\omega$-axis), then G can be expressed as

$$G = G_s + G_u$$

where G_s is stable and G_u is completely unstable.

This can be accomplished using Schur decomposition with eigenvalue ordering [L1,V1] and solving a set of linear equations (see Step 4 in Section 5.5(A)). Then, combining the balanced realizations of G_s and G_u leads to a minimal realization of G. Since the state matrix of either G_s or G_u is already in the upper real Schur form, the Lyapunov equations associated with gramians can be solved almost immediately [B2,G3].

This technique can be implemented very efficiently and reliably. It does not increase the complexity of the software since the balanced realization is essential to solve the best approximation problem (Section 5.4). The testing results are very satisfactory.

6.3 Model Reduction in H_∞ Synthesis

The importance of model reduction in control system design has long been recognized. For practical implementation, it is desired that the order of the controller can be reduced in a way such that the controlled system still satisfies the performance requirements. Typically, there are two ways to obtain a lower order controller: reducing the complexity of the plant model and using model reduction in the design process [E1,G2]. This section considers the latter issue.

High-order optimal controllers are usually derived when using H_∞ optimization. This can be seen clearly from the state-space formula shown in Section 5.5. Therefore, model reduction is inevitable from a practical point of view. Recall that the order of the controller $K = F_l(K_o, Q)$ is generically equal to the order of K_o, which is the same as the original interconnection structure P, plus the order of Q. Thus a natural first step in obtaining reduced-order controllers is to consider techniques that result in lower-order Q's. Another candidate for model reduction is the R in the **GDP**, since it is the complexity of R that affects not only the order of Q but also the computational burden involved in computing Q. The following analysis shows how the model reduction can be performed in the **GDP** with simple L_∞-norm bounds on the resulting loss of performance.

Assume that $\dot{Q}_{opt}$ is the optimal solution of the **GDP**:

$$\gamma_o = \min_{\dot{Q} \in RH_\infty} \left| \left| R - \begin{bmatrix} \dot{Q} & 0 \\ 0 & 0 \end{bmatrix} \right| \right|_\infty .$$

hen for model reduction, one has the following two results.

(i) Model reduction on R :

Suppose $\bar{R}$ is a reduced-order model of R, and $\bar{Q}_{opt}$ is the optimal solution of

$$\min_{\bar{Q} \in RH_\infty} \left| \left| \bar{R} - \begin{bmatrix} \bar{Q} & 0 \\ 0 & 0 \end{bmatrix} \right| \right|_\infty .$$

Define

$$\bar{\gamma}_{\circ} = \left|\left| R - \begin{bmatrix} \bar{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} .$$

The question is how much error $(\bar{\gamma}_{\circ} - \gamma_{\circ})$ is incurred if the reduced order model $\bar{R}$ is used in the **GDP**. This is found as follows:

$$\gamma_{\circ} \le \left|\left| R - \begin{bmatrix} \bar{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} = \left|\left| \bar{R} - \begin{bmatrix} \bar{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} + (R - \bar{R}) \right|\right|_{\infty}$$

$$\le \left|\left| \bar{R} - \begin{bmatrix} \bar{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} + ||R - \bar{R}||_{\infty}$$

$$\le \left|\left| \bar{R} - \begin{bmatrix} \hat{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} + ||R - \bar{R}||_{\infty}$$

$$= \left|\left| R - \begin{bmatrix} \hat{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} + (\bar{R} - R) \right|\right|_{\infty} + ||R - \bar{R}||_{\infty}$$

$$\le \left|\left| R - \begin{bmatrix} \hat{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} + 2||R - \bar{R}||_{\infty}$$

$$= \gamma_{\circ} + 2||R - R||_{\infty}$$

$$\therefore \quad \bar{\gamma}_{\circ} - \gamma_{\circ} \le 2||R - \bar{R}||_{\infty} \tag{6-1}$$

This inequality shows that the error is no more than $2||R - \bar{R}||_{\infty}$.

(ii) Model reduction on $\hat{Q}_{opt}$:

Suppose that $\hat{Q}_{app}$ is a reduced order model of the optimal solution $\hat{Q}_{opt}$. Then

$$\gamma_{\circ} \le \left|\left| R - \begin{bmatrix} \hat{Q}_{app} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty}$$

$$= \left|\left| R - \begin{bmatrix} \hat{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \hat{Q}_{opt} - \hat{Q}_{app} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty}$$

$$\le \left|\left| R - \begin{bmatrix} \hat{Q}_{opt} & 0 \\ 0 & 0 \end{bmatrix} \right|\right|_{\infty} + ||\hat{Q}_{opt} - \hat{Q}_{app}||_{\infty}$$

$$\le \gamma_{\circ} + ||\hat{Q}_{opt} - \hat{Q}_{app}||_{\infty} \tag{6-2}$$

Therefore, model reduction on $\hat{Q}_{opt}$ will introduce an error of no more than $||\hat{Q}_{opt} - \hat{Q}_{app}||_\infty$.

Suppose that model reduction in the H_∞ synthesis is done by the two steps:

(a) find the reduced-order model $\tilde{R}$ and the solution, $\tilde{Q}_{opt}$ of the corresponding GDP,

(b) find the reduced-order model, $\tilde{Q}_{app}$, of $\tilde{Q}_{opt}$.

Then,

$$\left|\left|R - \begin{bmatrix} \tilde{Q}_{app} & 0 \\ 0 & 0 \end{bmatrix}\right|\right|_\infty \leq 2||R - \tilde{R}||_\infty + ||\tilde{Q}_{opt} - \tilde{Q}_{app}||_\infty \tag{6-3}$$

This error bound can be derived easily by combining the results of (i) and (ii) above.

The above result is very encouraging since if the error bound in Eq. 6-3 is guaranteed to be small in model reduction, it will not affect the performance too much. Using either the method of truncation of the balanced realization [E1,G2] or the method of Hankel-norm approximation [G2], the reduced-order model can be found using reliable algorithms. Furthermore, both methods give the error bounds in terms of the L_∞-norm which are computable from the the second order modes of the given system (Section 5.4). A more detailed treatment on this subject can be found elsewhere [E1,G2]. Experience to date has shown that in many practical problems, both the order of $\tilde{R}$ and $\tilde{Q}_{opt}$ can be reduced significantly without incurring too much error.

APPENDIX A
LEMMAS

This appendix includes lemmas which were used in proving some of theorems.

Lemma A.1

Assume A, $B \in L_\infty$, then

$$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_\infty \leq \left\| \begin{bmatrix} \| A \|_\infty \\ \| B \|_\infty \end{bmatrix} \right\|_2 \leq \sqrt{2} \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_\infty .$$

Proof

(i) Left inequality:

By definition,

$$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_\infty = \sup_\omega \bar{\sigma} \begin{bmatrix} A \\ B \end{bmatrix}$$

$$= \sup_\omega \left[\lambda_{\max}(A^*A + B^*B) \right]^{1/2}$$

$$\leq \sup_\omega \left[\bar{\sigma}^2(A) + \bar{\sigma}^2(B) \right]^{1/2}$$

$$\leq \left[\| A \|_\infty^2 + \| B \|_\infty^2 \right]^{1/2}$$

$$= \left\| \begin{bmatrix} \| A \|_\infty \\ \| B \|_\infty \end{bmatrix} \right\|_2$$

(ii) Right inequality:

Since

$$\| A \|_\infty \leq \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_\infty \quad \text{and} \quad \| B \|_\infty \leq \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_\infty$$

$$\implies \quad \frac{1}{\left\|\left\|\begin{bmatrix} A \\ B \end{bmatrix}\right\|\right\|_\infty} \left\|\left\|\begin{bmatrix} \|A\|_\infty \\ \|B\|_\infty \end{bmatrix}\right\|\right\|_2 \leq \left\|\left\|\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\|\right\|_2 = \sqrt{2}$$

$$\implies \quad \left\|\left\|\begin{bmatrix} \|A\|_\infty \\ \|B\|_\infty \end{bmatrix}\right\|\right\|_2 \leq \sqrt{2} \left\|\left\|\begin{bmatrix} A \\ B \end{bmatrix}\right\|\right\|_\infty .$$

<div align="right">QED</div>

Lemma A.2

$$\left\|\left\|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right\|\right\|_\infty \leq \left\|\left\|\begin{bmatrix} \|A\|_\infty & \|B\|_\infty \\ \|C\|_\infty & \|D\|_\infty \end{bmatrix}\right\|\right\|_F \leq 2 \left\|\left\|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right\|\right\|_\infty .$$

Proof

(i) Left inequality:

From Lemma A.1,

$$\left\|\left\|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right\|\right\|_\infty \leq \left\|\left\|\begin{bmatrix} \|[A \quad C]\|_\infty \\ \|[B \quad D]\|_\infty \end{bmatrix}\right\|\right\|_2$$

$$= \left[\|[A \quad C]\|_\infty^2 + \|[B \quad D]\|_\infty^2 \right]^{1/2}$$

$$\leq \left[\|\|[\|A\|_\infty \quad \|C\|_\infty]\|\|_2^2 + \|\|[\|B\|_\infty \quad \|D\|_\infty]\|\|_2^2 \right]^{1/2}$$

$$= \left[\|A\|_\infty^2 + \|B\|_\infty^2 + \|C\|_\infty^2 + \|D\|_\infty^2 \right]^{1/2}$$

$$= \left\|\left\|\begin{bmatrix} \|A\|_\infty & \|B\|_\infty \\ \|C\|_\infty & \|D\|_\infty \end{bmatrix}\right\|\right\|_F .$$

(ii) Right inequality:

$$\left\|\left\|\begin{bmatrix} \|A\|_\infty & \|B\|_\infty \\ \|C\|_\infty & \|D\|_\infty \end{bmatrix}\right\|\right\|_F = \left[\|\|[\|A\|_\infty \quad \|C\|_\infty]\|\|_2^2 + \|\|[\|B\|_\infty \quad \|D\|_\infty]\|\|_2^2 \right]^{1/2}$$

$$\leq \left[2 \|[A \quad C]\|_\infty^2 + 2 \|[B \quad D]\|_\infty^2 \right]^{1/2}$$

$$= \sqrt{2} \left[\left\| \left\| \begin{bmatrix} A & C \end{bmatrix} \right\| \right\|_\infty^2 + \left\| \left\| \begin{bmatrix} B & D \end{bmatrix} \right\| \right\|_\infty^2 \right]^{1/2}$$

$$= \sqrt{2} \left\| \left\| \begin{bmatrix} \left\| \begin{bmatrix} A & C \end{bmatrix} \right\|_\infty \\ \left\| \begin{bmatrix} B & D \end{bmatrix} \right\|_\infty \end{bmatrix} \right\| \right\|_2$$

$$\leq \sqrt{2} \cdot \sqrt{2} \left\| \left\| \begin{bmatrix} A & B \\ C & D \end{bmatrix} \right\| \right\|_\infty$$

$$= 2 \left\| \left\| \begin{bmatrix} A & B \\ C & D \end{bmatrix} \right\| \right\|_\infty .$$

<div align="right">QED</div>

Lemma A.3

Assume $A = \begin{bmatrix} X & Y \end{bmatrix} \in L_\infty$. Then,

$$T^*T = \gamma^2 I - A^*A$$

where

$$T = \begin{bmatrix} \gamma(I - \tilde{L}^*\tilde{L})^{1/2} & 0 \\ -(S^{-1})^*Y^*X & S \end{bmatrix}$$

$$S = (\gamma^2 I - Y^*Y)^{1/2}$$

and

$$\tilde{L} = (\gamma^2 I - YY^*)^{-1/2}X$$

Proof

Instead of verifying the equality directly by forming the product T^*T which is just the routine algebra, a different but more informative approach will be used here.

Recall that in the constant positive semi-definite matrix case, a Cholesky factorization can be found with either lower or upper triangular factors. This fact can be extended to the "block" triangular form. Therefore, the matrix T can be assumed to have the following block lower triangular form:

$$T = \begin{bmatrix} W & 0 \\ V & S \end{bmatrix} .$$

Then

$$T^*T = \gamma^2 I - A^*A$$

$$\Longrightarrow \quad \begin{bmatrix} W^*W + V^*V & V^*S \\ S^*V & S^*S \end{bmatrix} = \begin{bmatrix} \gamma^2 I - X^*X & -X^*Y \\ -Y^*X & \gamma^2 I - Y^*Y \end{bmatrix}$$

$$\therefore \quad W^*W + V^*V = \gamma^2 I - Y^*Y$$

$$S^*V = -Y^*X$$

$$S^*S = \gamma^2 I - Y^*Y$$

Therefore,

$$S = (\gamma^2 I - Y^*Y)^{1/2}$$

$$V = -(S^*)^{-1}Y^*X$$

and

$$\begin{aligned}
W^*W &= \gamma^2 I - X^*X - V^*V \\
&= \gamma^2 I - X^*X - X^*Y(S^{-1})(S^{-1})^*Y^*X \\
&= \gamma^2 I - X^*[I + Y(S^*S)^{-1}Y^*]X \\
&= \gamma^2 I - X^*[I + Y(\gamma^2 I - Y^*Y)^{-1}Y^*]X \\
&= \gamma^2 I - X^*[I + (\gamma^2 I - YY^*)^{-1}YY^*]X \\
&= \gamma^2 I - \gamma^2 X^*(\gamma^2 I - YY^*)^{-1}X \\
&= \gamma^2(I - \tilde{L}^*\tilde{L})
\end{aligned}$$

$$\Longrightarrow \quad W = \gamma(I - \tilde{L}^*\tilde{L})^{1/2}$$

QED

The following two lemmas are well-known.

Lemma A.4 (Solution of Sylvester Equations)

Consider the Sylvester equation

$$AX + XB = C \qquad \text{(A-1)}$$

where $A \in \mathbf{R}^{n \times n}$, $B \in \mathbf{R}^{m \times m}$, and $C \in \mathbf{R}^{n \times m}$ are given matrices. Then, there exists a unique solution $X \in \mathbf{R}^{n \times m}$ if and only if

$$\text{Re}[\lambda_i(A) + \lambda_j(B)] \neq 0, \qquad \forall \ i = 1,\ldots\ldots,n \ \ and \ \ j = 1,\ldots\ldots,m \ .$$

Remark

In particular, if $B = A^T$, Eq. A-1 is called the "Lyapunov Equation" and the necessary and sufficient condition for the existence of a unique solution will be that

$$\text{Re}[\lambda_i(A) + \lambda_j(A)] \neq 0, \ \forall \ i , j = 1,\ldots\ldots,n \ .$$

Lemma A.5 (Solution of Linear Equations)

Consider the linear equation

$$A \ X = B$$

where $A \in \mathbf{R}^{n \times n}$ and $B \in \mathbf{R}^{n \times m}$ are given matrices.

The following statements are equivalent:

(i)   there exists a solution $X \in \mathbf{R}^{n \times m}$.

(ii)   the columns of $B \in Range \ (A)$.

(iii) $rank \begin{bmatrix} A & B \end{bmatrix} = rank \begin{bmatrix} A \end{bmatrix}$.

(iv) $\mathbf{Ker} \ (A^T) \subset \mathbf{Ker} \ (B^T)$.

APPENDIX B
ISOMORPHISM BETWEEN
THE HALF-PLANE AND THE UNIT DISC

Define the transformation

$$z = \frac{s-1}{s+1} \qquad (s = \frac{1+z}{1-z}) \tag{B-1}$$

which maps the right half plane ($\mathrm{Re}(s) \geq 0$) onto the unit disc ($|z| \leq 1$). Therefore, the relation between a point $j\omega$ on the imaginary axis and the corresponding point $e^{j\theta}$ on the unit circle is, from Eq. B-1,

$$e^{j\theta} = \frac{j\omega-1}{j\omega+1}.$$

Also define the function

$$\psi(s) = \frac{\sqrt{2}}{(s+1)}.$$

This yields $d\theta = -\dfrac{2}{\omega^2+1} d\omega = -|\psi(j\omega)|^2 d\omega$.

Let T be the unit circle in z-domain and $j\mathbb{R}$ be the imaginary axis in s-plane. This implies that the mapping

$$\bar{f} \rightarrow \psi f \ : H_2(T) \rightarrow H_2(j\mathbb{R}),$$

where $\bar{f}(z) = f(s)\Big|_{s=\frac{1+z}{1-z}}$, is an isomorphism. Similarly,

$$\bar{f} \rightarrow \psi f \ : H_2(T)^{\perp} \rightarrow H_2(j\mathbb{R})^{\perp}$$

is an isomorphism; note that if $\bar{f} \in H_2(T)^{\perp}$, then $\bar{f} = 0$ at $z = \infty$, so that $f = 0$ at $s = -1$, and hence ψf is analytic in $\mathrm{Re}\ s < 0$.

APPENDIX C

OPERATIONS ON LINEAR SYSTEMS

This appendix describes some of standard operations on linear systems in terms of transfer functions and their realizations which were used in previous chapters.

1. Cascade

$$G_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] , \qquad G_2 = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] .$$

$$
\begin{aligned}
G_1 G_2 &= \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] * \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] \\
&= \left[\begin{array}{cc|c} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1 C_2 & D_1 D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1 C_2 & A_1 & B_1 D_2 \\ \hline D_1 C_2 & C_1 & D_1 D_2 \end{array}\right] .
\end{aligned}
$$

Note: This realization may not be minimal.

2. Change of Variables

$$x \rightarrow \dot{x} = Tx$$

$$y \rightarrow \dot{y} = Ry$$

$$u \rightarrow \dot{u} = Pu$$

$$
\begin{aligned}
\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \rightarrow \left[\begin{array}{c|c} \dot{A} & \dot{B} \\ \hline \dot{C} & \dot{D} \end{array}\right] &= \left[\begin{array}{cc} T & 0 \\ 0 & R \end{array}\right] \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \left[\begin{array}{cc} T^{-1} & 0 \\ 0 & P^{-1} \end{array}\right] \\
&= \left[\begin{array}{c|c} TAT^{-1} & TBP^{-1} \\ \hline RCT^{-1} & RDP^{-1} \end{array}\right] .
\end{aligned}
$$

3. State Feedback

$$u \rightarrow \dot{u} + Fz$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \rightarrow \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\begin{bmatrix} I & 0 \\ F & I \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline C+DF & D \end{array}\right].$$

4. Output Injection

$$\dot{z} = Az + Bu \rightarrow \dot{z} = A\dot{z} + Bu + Hy$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \rightarrow \begin{bmatrix} I & H \\ 0 & I \end{bmatrix}\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{c|c} A+HC & B+HD \\ \hline C & D \end{array}\right].$$

5. Transpose (Dual)

$$G \rightarrow G^T$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \rightarrow \left[\begin{array}{c|c} A^T & C^T \\ \hline B^T & D^T \end{array}\right].$$

6. Conjugate

$$G \rightarrow G^*$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \rightarrow \left[\begin{array}{c|c} -A^T & -C^T \\ \hline B^T & D^T \end{array}\right] \quad \left(\text{or } \left[\begin{array}{c|c} -A^T & C^T \\ \hline -B^T & D^T \end{array}\right]\right).$$

REFERENCES

[A1]   V.M. Adamjan, D.Z. Arov, and M.G. Krein, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem," *Math. USSR Sbornik*, 15, No. 1, 1971, pp. 31-73.

[A2]   V.M. Adamjan, D.Z. Arov, and M.G. Krein, "Infinite Hankel block matrices and related extension problems," *AMS Transl.*, 111, 1978, pp. 133-156.

[A3]   B.D.O. Anderson, "An algebraic solution to the spectral factorization problem," *IEEE Trans. Auto. Control*, 12, 1967, pp. 410-414.

[A4]   B.D.O. Anderson and S. Vongpanitlerd, *"Network Analysis and Synthesis: A Modern Systems Theory Approach"*, Prentice-Hall, 1973.

[B1]   J.A. Ball and J.W. Helton, "A Beurling-Lax theorem for the Lie group $U(m,n)$ which contains most classical interpolation theory," *J. Operator Theory*, 9, 1983, pp. 107-142.

[B2]   R.H. Bartels and G.W. Stewart, "A solution of the equation $AX+XB=C$," *Commun. ACM*, 15, 1972, pp. 820-826.

[B3]   M. Bettayeb, L.M. Silverman, and M.G. Safonov, "Optimal approximation of continuous-time systems," *Proc. IEEE Conf. on Deci. and Control*, Albuquerque, NM, 1980.

[B4]   H.W. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand: Princeton, 1945.

[B5]   A. Brown and P.R. Halmos, "Algebraic properties of Toeplitz operators," *J. Reine Angew Math.*, 213, 1964, pp. 89-102.

[C1]   B.C. Chang and J.B. Pearson, "Optimal disturbance reduction in linear multivariable systems," *IEEE Trans. Auto. Control*, 29, Oct. 1984, pp. 880-887.

[C2]   C.C. Chu and J.C. Doyle, "On inner-outer and spectral factorizations," *Proc. IEEE Conf. on Deci. and Control*, Las Vegas, NV, 1984.

[C3]   C.C. Chu and J.C. Doyle, "Computational issues in μ-synthesis," *IEEE 2nd CASCAD*, Santa Barbara, CA, March 1985.

[C4]   C.C. Chu and J.C. Doyle, "The general distance problem in H_∞ synthesis," *Proc. IEEE Conf. on Deci. and Control*, Fort Lauderdale, FL, 1985.

[C5]   W.A. Coppel, "Matrix Quadratic Equation," *Bull. Austral. Math. Soc.*, 10, 1974, pp. 377-401.

[D1]   C. Davis, W.M. Kahan and H.F. Weinberger, "Norm-preserving dilations and their applications to optimal error bounds," *SIAM J. Numer. Anal.*, 19, 1982, pp. 445-469.

[D2]   P.H. Delsarte, Y. Genin, and Y. Kamp, "The Nevanlinna-Pick problem for matrix-valued functions," *SIAM J. Appl. Math.*, 36, Feb. 1979, pp. 47-61.

[D3]   C.A. Descer and M. Vidyasagar, *"Feedback Systems: Input-Output Properties"*, New York: Academic Press, 1975.

[D4]   C.A. Desoer and W.S. Chan, "The feedback interconnection of linear time-invariant systems," *J. Franklin Inst.*, 300, 1975, pp. 335-351.

[D5]   C.A. Desoer, R.W. Liu, J. Murray, and R. Saeks, "Feedback system design: The fractional representation approach to analysis and synthesis," *IEEE Trans. Auto. Control*, 25, June 1980, pp. 399-412.

[D6]   J.J. Dongarra, C.B. Moler, J.R. Bunch, and G.W. Stewart, *"LINPACK User's Guide"*, Philadelphia, PA: SIAM, 1979.

[D7]   R.G. Douglas, *"Banach Algebra Techniques in Operator Theory"*, Academic Press, New York and London, 1972.

[D8]   R.G. Douglas, *"Banach algebra techniques in the theory of Toeplitz operators "*, C.B.M.S. Regional Conference, No.15, A.M.S., Providence, R.I., 1973.

[D9]   R.G. Douglas, "Local Toeplitz Operators," *Proc. London Math. Soc.*, 36, 1978, pp. 243-272.

[D10]   J.C. Doyle and G. Stein, "Singular values and feedback: design examples," 1978 Allerton Conference on Communications, Control, and Computing.

[D11]   J.C. Doyle, "Robustness of multiloop linear feedback systems," *Proc. IEEE Conf. on Deci. and Control*, San Diego, CA, 1978.

[D12]   J.C. Doyle and G. Stein, "Multivariable feedback design: concepts for a classical/modern synthesis," *IEEE Trans. Auto. Control*, 26, Feb. 1981, pp. 4-16.

[D13]   J.C. Doyle, "Analysis of feedback systems with structured uncertainty," *IEE Proceedings, Part D*, 129, No. 6, Nov. 1982.

[D14]   J.C. Doyle, J.E. Wall, and G. Stein, "Performance and robustness analysis for structured uncertainty," *Proc. IEEE Conf. on Deci. and Control*, Orlando, FL, 1982.

[D15]   J.C. Doyle, "Synthesis of robust controllers and filters," *Proc. IEEE Conf. on Deci. and Control*, San Antonio, TX, 1983.

[D16]   J.C. Doyle, *Lecture Notes*, 1984 ONR/Honeywell Workshop on Advances in Multivariable Control, October 8-10, 1984, Minneapolis, MN.

[D17]   J.C. Doyle and C.C. Chu, "Matrix interpolation and H_∞ performance bounds," *Proc. Amer. Contr. Conf.*, Boston, MA, 1985.

[D18]   P.L. Duren, *"The Theory of H^p-spaces"*, New York: Academic Press, 1970.

[D18]   M.A. Dahleh and J.B. Pearson, " l^1 - Optimal Feedback Controllers for MIMO Discrete-Time Systems", Rice University EE Dept. Tech. Rept. TR 8602, February 1986.

[E1]   D.F. Enns, "Model reduction for control system design," Ph.D. Dissertation, Stanford University, 1984.

[F1]   A. Feintuch and B.A. Francis, "Uniformly optimal control of linear time-varying systems," *Systems & Control Letters*, 5, Oct. 1984, pp. 67-71.

[F2]   A. Feintuch and B.A. Francis, "Uniformly optimal control of linear feedback systems," (to appear in *Automatica*).

[F3]   Y.K. Foo and I. Postlethwaite, "An H^∞-minimax approach to the design of robust control systems," *Systems & Control Letters*, 5, Nov. 1984, pp. 81-88.

[F4]   B.A. Francis and G. Zames, "On H^∞- optimal sensitivity theory for SISO feedback systems," *IEEE Trans. Auto. Control*, 29, 1984, pp. 9-16.

[F5]   B.A. Francis, J.W. Helton, and G. Zames, "H^∞-optimal feedback controllers for linear multivariable systems," *IEEE Trans. Auto. Control*, 29, Oct. 1984, pp. 888-900.

[F6]   B.A. Francis, private communication.

[F7]   Francis, B.A., 1983, Lecture notes presented at Linkoping University.

[F8]   B.A. Francis, and J.C. Doyle, 1985, University of Toronto Systems Control Group Report #8501, also to appear, *SIAM J. Cntrl. and Opt.*

[F8]   M.K.H. Fan, and A.L.T. Tits, "A New Formula for the Structured Singular Value," 1985 IEEE CDC

[G1]   B.S. Garbow, J.M.Boyle, J.J. Dongarra, and J.B. Moler, "*Matrix Eigensystem Routines - EISPACK Guide Extension*", New York: Spring-Verlag, 1977.

[G2]   K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds," *Int. J. Control*, 39, No. 6, 1984, pp. 1115-1193.

[G3]   G.H. Golub, S. Nash, and C. Van Loan, "A Hessenberg-Schur method for the problem $AX+XB=C$," *IEEE Trans. Auto. Control*, 24, Dec. 1979, pp. 909-913.

[G4]   Gohberg , I. and Krein, M.G., 1969, *Introduction to the theory of linear nonselfadjoint operators*, Translations of Mathematical Monographs, Vol.18, AMS.

[H1]   S.J. Hammarling, "Numerical solution of the stable, non-negative definite Lyapunov equation," *IMA J. Numer. Anal.*, 2, 1982, pp. 303-323.

[H2]   J.W. Helton, "Broadbanding: gain equalization directly from data," *IEEE Trans. Circuits and Sys.*, 28, Dec. 1981, pp. 1125-1137.

[H3]   J.W. Helton, "Worst case analysis in the frequency domain: the H^∞ approach to control," *Proc. IEEE Conf. on Deci. and Control*, San Antonio, TX, 1983.

[H4]   I.M. Horowitz, "*Synthesis of Feedback Systems*", New York: Academic Press, 1963.

[J1]   Jonckheere, E. and Silverman, L. M., 1980, *IEEE Trans. Auto. Control*, AC-25, pp. 880-888.

[J2]   Jonckheere, E. and Verma, M., submitted 1985, *Systems & Control Letters*.

[K1]   T. Kailath, "*Linear Systems*", Englewood Cliffs, NJ: Prentice-Hall, 1980.

[K2]   R.E. Kalman and R.S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, Ser. D, J. Basic Eng.*, 83, Dec. 1961, pp. 95-107.

[K3]   T. Kato, "*Perturbation Theory for Linear Operators*", Springer-Verlag, 1976.

[K4]   P.P. Khargonekar and E.D. Sontag, "On the relation between stable matrix fraction factorizations and regulable realizations of linear systems over rings," *IEEE Trans. Auto. Control*, 27, June 1982, pp. 627-638.

[K5]   P. Koosis, "*Introduction to H^p-spaces*", Cambridge: Cambridge University Press, 1980.

[K6]   V. Kucera, "A contribution to matrix quadratic equation," *IEEE Trans. Auto. Control*, 17, 1972, pp. 344-347.

[K7]   S.Y. Kung and D.W. Lin, "Optimal Hankel-norm model reduction: multivariable systems," *IEEE Trans. Auto. Control*, 26, 1981, pp. 832-852.

[K8]   H. Kwakernaak and R. Sivan, "*Linear Optimal Control Systems*", New York: Wiley-Interscience, 1972.

[K9]   H. Kwakernaak, "Robustness optimization of linear feedback systems," *Proc. IEEE Conf. on Deci. and Control*, San Antonio, TX, 1983.

[K10]   H. Kwakernaak, "Minimax frequency domain optimization of multivariable linear feedback systems," *IFAC World Congress*, Budapest, Hungary, 1984.

[L1]   A.J. Laub, "A Schur method for solving algebraic Riccati equations," *IEEE Trans. Auto. Control*, 24, Dec. 1979, pp. 913-921.

[L2]   A.J. Laub, "Computation of balancing transformations," *Proc. Joint Auto. Contr. Conf.*, San Francisco, CA, 1980.

[L3]   "Special Issue on Linear-Quadratic-Gaussian Problem," *IEEE Trans. Auto. Control*, 16, Dec. 1971.

[L4]   D.G. Luenberger, "Observers for multivariable systems," *IEEE Trans. Auto. Control*, 11, 1966, pp. 190-197.

[M1]   C.F. Matensson, "On the matrix Riccati equation," *Info. Sci.*, 3, 1971, pp. 17-49.

[M2]   B.P. Molinari, "The stabilizing solution of the algebraic Riccati equation," *SIAM J. Control*, 11, 1973, pp. 262-271.

[M3]   B.P. Molinari, "Equivalence relations for the algebraic Riccati equation," *SIAM J. Control*, 11, 1973, pp. 272-285.

[M4]   B.C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Trans. Auto. Control*, 26, Feb. 1981.

[N1]   Z. Nehari, "On bounded bilinear forms," *Ann. of Math*, 65, 1957, pp. 153-162.

[N2]   C.N. Nett, C.A. Jacobson, and M.J. Balas, "A connection between state-space and doubly coprime fractional representations," *IEEE Trans. Auto. Control*, 29, Sept. 1984, pp. 831-832.

[P1]   S. Parrott, "On a quotient norm and the Sz-Nagy Foias lifting theorem," *J. of Functional Analysis*, 30, 1978, pp. 311-328.

[P2]   L. Pernebo and L.M. Silverman, "Model reduction via balanced state-space representations," *IEEE Trans. Auto. Control*, 27, April 1982, pp. 382-387.

[P3]   J.E. Potter, "Matrix quadratic solutions," *SIAM J. Control*, 14, 1966, pp. 496-501.

[P4]   S.C. Power, "*Hankel operators on Hilbert space*", Pitman Publishing Inc. 1982.

[S1]   M.G. Safonov and M. Verma, "Multivariable L^∞-sensitivity optimization and Hankel approximation," *Proc. Amer. Contr. Conf.*, San Francisio, CA, 1983.

[S2]   D.E. Sarason, "Generalized interpolation in H^∞," *Trans. Amer. Math. Soc.*, 127, May 1967, pp. 179-203.

[S3]   D.E. Sarason, "On products of Toeplitz operators," *Acta Sci. Math. (Szeged)*, 35, 1973, pp. 7-12.

[S4]   L.M. Silverman and M. Bettayeb, "Optimal approximation of linear systems," *Proc. Joint Auto. Contr. Conf.*, San Francisco, CA, 1980.

[S5]   M.G. Safonov and J.C. Doyle, "Conservativeness of Robustness Singular Values," in *Multivariable Control* (ed. S.G. Tzafestas). Dordrecht, Holland: D. Riedel, 1984.

[S5]   M.G. Safonov, "Optimal Diagonal Scaling for Infinity Norm Optimization," 1985 ACC, Boston.

[V1]   P. Van Dooren, "A generalized eigenvalue approach for solving Riccati equations," *SIAM J. Sci. Stat. Comput.*, 2, June 1981, pp. 121-135.

[V2]   M. Vidyasager, H. Schneider and B.A. Francis, "Algebraic and topological aspects of feedback stabilization," *IEEE Trans. Auto. Control*, 27, 1982, pp. 880-894.

[V3]   M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, MA: MIT Press, 1985.

[V4]   Verma, M. and Jonckheere, E., 1984, *Systems & Control Letters*, 4, pp. 125-129.

[W1]   J.C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Auto. Control*, 16, 1971, pp. 621-634.

[Y1]   D.C. Youla, "On the factorization of rational matrices," *IRE Trans. Information Theory*, IT-7, 1961, pp. 172-189.

[Y2]   D.C. Youla, H.A. Jabr, and J.J. Bongiorno, Jr., "Modern Wiener-Hopf design of optimal controllers, Part II: The multivariable case," *IEEE Trans. Auto. Control*, 21, June 1976, pp. 319-338.

[Z1]   G. Zames, "On the input-output stability of nonlinear time-varying feedback systems, Pt. I and II," *IEEE Trans. Auto. Control*, 11, April 1966, pp. 228-238; July 1966, pp. 465-477.

[Z2]   G. Zames, "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Auto. Control*, 26, April 1981, pp. 301-320.

[Z3]   G. Zames and B.A. Francis, "Feedback, minimax sensitivity, and optimal robustness," *IEEE Trans. Auto. Control*, 28, 1983, pp. 585-601.

END

1 — 81

DTIC